
**Edinburgh School of Economics**

# Discussion Paper Series

Number 94

# *Matching and Sorting with Horizontal Heterogeneity*

**Simon Clark** (University of Edinburgh)

Date
September 2003

***Published by***
School of Economics
University of Edinburgh
30 -31 Buccleuch Place
Edinburgh EH8 9JT
+44 (0)131 650 8361
http://www.ed.ac.uk/schools-departments/economics

THE UNIVERSITY *of* EDINBURGH

# Matching and Sorting with Horizontal Heterogeneity

Simon Clark[1]

September 2003

ABSTRACT: This paper examine a class of two-sided matching problems with non-transferable utility. Agents are horizontally differentiated, and each would prefer to be matched with a similar partner; in short, "like attracts like". Although such preferences imply a unique stable matching, the degree of assortment in equilibrium is found to depend critically on the distribution of characteristics among the two sexes. In particular, the greater the difference between men and women, the greater the tendency to negative assortment. Constraints on who can match with whom may improve welfare and we interpret this as a theory of social stratification.

KEYWORDS: Matching; sorting; uniqueness; horizontal heterogeneity; marriage market; social stratification.

*JEL* CLASSIFICATION NUMBER: C7

[1] Department of Economics, Edinburgh University, William Robertson Building, 55 George Square, Edinburgh EH8 9JY, UK.

# 1 INTRODUCTION

Two-sided matching has been described as one of the great sucesses of game theory (Aumann, 1990). The matching framework provides a natural and fruitful way to model a wide variety of problems where the side of a transaction to which each agent belongs can be taken as given; e.g. marriage (Gale and Shapley, 1962; Becker, 1981), university admissions (Gale and Shapley, 1962), hospital intern programs (Roth, 1984), and auctions (Roth and Sotomayor, 1990). Furthermore, a standard matching set-up often sits at the centre of more elaborate models in which the primary focus is search or the role of matching frictions, for example in labour markets (Coles and Burdett, 1999), in bilateral exchange (Kiyotaki and Wright, 1989), or in the rationing of electicity supply (McAfee, 2002).

An obvious question to ask of the outcome of many matching markets is whether they display positive assortment. For instance, do richer men marry richer women? Or do cleverer students go to more prestigious universities? In the case of fully transferable utility, a well established result is that regardless of the distribution of types matching will display positive assortment if the combined output of two matched agents is a supermodular function of their characteristics.[2] If utility is partially but not fully transferable (i.e. if the utility possibility frontier for any pair of agents is downward sloping but not linear) then, as Legros and Newman (2003) show, positive assortment can be guaranteed by a further condition on how the degree of transferability depends on the two agents' types.

If utility is not transferable at all (the utility possibility frontier for any pair of agents is rectangular), matters are very different. Agents can assess and rank potential partners without the need to take into account where they will be on the utility possibility frontier. When considering beauty, or intelligence, or prestige, it might be reasonable to assume that all agents on one side of the market agree on how they rank the agents on the other side: everyone would prefer a more attractive spouse; all students would prefer to go to a more prestigious university; all universities would prefer to admit a cleverer student. With agents displaying such *vertical* heterogeneity, the most desirable agent on one side will match with the most desirable on the other side, the second most desirable agents will match with each other, and so on. Thus positive assortment will arise if there are no

[2] Recently, Legros and Newman (2002) have provided weaker sufficent conditions for positive assortment.

matching frictions and agents can freely choose with whom they match.

On the other hand, agents may be *horizontally* differentiated, and prefer to match with someone who is similar to them, or who fits in with their own objectives or capabilities; in brief, like may be attracted to like. For example, a man might prefer a wife of the same height, or who has similar tastes. A student might prefer to go to a college where the courses are pitched at a level suitable to his or her ability; a college might prefer a less able student because it has a mission, and the specialist expertise, to educate the less gifted, not just the cleverest. A hospital specialising in cancer care and research may prefer the academically oriented intern who has done well in oncology, whereas the hospital with a vacancy in its busy city centre Accident and Emergency department would prefer an intern with practical skills who can work under pressure. With preferences such as these, it is not at all clear that positive sorting will be the outcome, however desirable. For example, it may be that for a hospital there are no graduating medical students with the exact mix of abilities and character that it seeks; but its preferred candidate, of those available, would prefer to work at another hospital. Similarly, for a prospective intern there may be no vacancy in her ideal hospital. Who then gets matched with whom is of course the very stuff of matching theory. But somewhat remarkably, once we move away from fully transferable utility, "for the analyst seeking to characterize the equilibrium matching patterns in such settings, there is little theoretical guidance" (Legros and Newman, 2003, p.2). This paper goes some way to fill that gap. I analyse matching and sorting when utility is non-transferable and like attracts like. However, rather than look for conditions under which we get complete positive assortment, I focus on what determines the degree of assortment.

To explain why some matching problems result in a greater degree of assortment than others we should ideally use an equilibrium concept that produces unique equilibria. The central solution concept in matching theory is the *stable matching*, which pairs the members of one set (e.g. men) with those of another, disjoint, set (e.g. women), in such a way that no man and woman who are not paired with each other would both prefer to leave their partners and marry each other.[3] Gale and Shapley proved in 1962 that such an equilibrium exists, but without restrictions on preferences there will generally be multiple stable matchings. With $n$ men and $n$ women, there are $n!$ possible matchings; if $n$ is a power of 2 it is possible to find preferences such that there are at least $2^{n-1}$ stable matchings (Irving and Leather,

[3]The set of the stable matchings can be seen as the core of a game in which the only allowable coalitions are pairs consisiting of one man and one woman.

1986).

To illustrate how multiple stable matchings can arise, and how uniqueness can result from restrictions on preferences, here is a "Greek tragedy", adapted from Racine's play *Andromaque*. Orestes, son of the Greek king Agamemnon, loves Helen of Troy's daughter Hermione, who loves Achilles' son Pyrrhus, who loves Andromaque, widow of the Trojan hero Hector. Departing from Racine, let us suppose that Andromaque loves Orestes. The term "loves" is to be interpreted as "prefers, out of the members of the opposite sex". Figure 1 illustrates. There are two possible matchings: either Orestes is paired with Hermione, and Pyrrhus with Andromaque, or Orestes is paired with Andromaque, and Pyrrhus with Hermione. Both of these matchings are stable.[4]

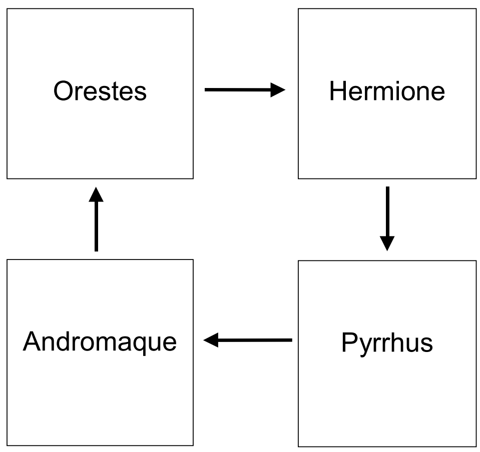


Figure 1: A Greek Tragedy. Arrows indicate the preferred member of the opposite sex.

As Figure 1 suggests, the preferences described above display a certain circularity, which must be broken to achieve uniqueness. Two ways suggest themselves. If Orestes and Pyrrhus both love Andromaque, then in equilibrium she would have to be matched with her preferred man, Orestes, leaving Pyrrhus paired with Hermione. Thus one approach is to assume all agents on one side of the market have the same preferences, with the actual matching detemined by the preferences of the other side (the most preferred women gets her man, the second most preferred women gets her preferred man from those who are left, and so on). This route to uniqueness was identified by Gusfield and Irving (1989), although their main interest was in algorithms designed to find all stable matches. A special case arises if both sides have

[4]I am assuming any marriage is better then remaining single.

common preferences, as in the cases of vertical differentiation described above.

Alternatively, our population of Greeks may be horizontally differentiated. Suppose that each person prefers someone closer to them in height, and consider two possibilities:

**Case 1:** Orestes and Hermione are both 1.70m tall, and Pyrrhus and Andromaque are both 1.80m tall. Then there is a unique stable matching, in which Orestes is matched with Hermione and Pyrrhus with Andromaque. This matching displays positive assortment, the taller man being paired with the taller woman.

**Case 2:** Hermione is 1.70m tall, Andromaque and Orestes are 1.80m tall, and Pyrrhus is 1.90m tall. Then Orestes and Andromaque, being the same height, are perfectly matched, and must be paired in equilibrium, leaving Hermione and Pyrrhus also paired in the unique stable matching. In Case 2, the taller man is paired with the shorter woman, so the matching displays negative assortment.

These two cases illustrate the themes that are developed in the remainder of this paper. Firstly, if preferences are such that each agent would most like to match with someone who is similar to themselves, then there is a unique stable matching. This is a straightforward application of a more general result concerning preferences that satisfy a *No Crossing Condition*, due to Clark (2002), and is discussed in Section 2. Secondly, uniqueness of the stable matching need not imply a positive association between partners' characteristics. Even in the case where everyone would prefer to match with a partner of a similar type, not everyone will get their preferred mate, as Hermione and Pyrrhus found out in the previous paragraph. A comparison of Cases 1 and 2 above suggests that the distribution of agents' characteristics is critical in determining the degree of assortment. Section 3 formalises this by taking the agents's characteristics to be evenly (and in the limit uniformly) distributed, and deriving an exact form of one measure of the degree of sorting, the rank correlation between agents' characteristics and their partners'. We then show how this measure depends on the parameters of the two distributions. Section 4 considers the welfare loss when sorting is not completely positive, and considers possible responses. We show that constraints on who can match with whom typically improve welfare and we tentatively interpret this as a theory of social stratification. Section 5 of the paper concludes.

# 2 WHEN LIKE ATTRACTS LIKE

## 2.1 The matching framework

We consider two finite and disjoint sets, both with $n$ elements. For simplicity we shall refer to these as a set of men $M = \{m_1, m_2, ...m_n\}$ and a set of women $W = \{w_1, w_2, ...w_n\}$, and consider them as ordered subsets of $R$, the real line, so $m_i < m_j$ and $w_i < w_j$ if $i < j$. We refer to $P = M \cup W$ as the population. We can thus identify any member of $P$ with a number, which for convenience we refer to as his or her height. In order to capture the idea that each person would prefer to be paired with a similar member of the opposite sex we assume that a man $m$ prefers woman $w_i$ to woman $w_j$ if $|m - w_i| < |m - w_j|$. Similarly, woman $w$ prefers man $m_i$ to man $m_j$ if $|w - m_i| < |w - m_j|$. The matching framework we adopt requires each agent to have a strict preference ranking over members of the opposite sex. Since $M$, $W \subset R$, no two men or no two women are of exactly the same height, but we further assume that if $m - w_i = w_j - m > 0$ then $m$ prefers $w_i$ to $w_j$; and if $w - m_i = m_j - w > 0$ then $w$ prefers $m_j$ to $m_i$.[5] Lastly, we assume that each person always prefers any partner of the opposite sex to remaining single.

## 2.2 Existence of equilibrium

Having described the marriage market, we now define the equilibrium concept.

**Definition 1** *A* matching $\mu$ *of the population* $P$ *is a one-to-one function from* $P$ *onto itself such that (i)* $m = \mu(w)$ *if and only if* $w = \mu(m)$; *(ii) if* $m \in M$ *then* $\mu(m) \in W$ *and if* $w \in W$ *then* $\mu(w) \in M$.

**Definition 2** *A matching* $\mu$ *can be* blocked *by a pair* $(m, w) \in M \times W$ *for whom* $m \neq \mu(w)$ *if* $m$ *strictly prefers* $w$ *to* $\mu(m)$ *and* $w$ *strictly prefers* $m$ *to* $\mu(w)$. *A matching* $\mu$ *is* stable *if it cannot be blocked by any pair.*

Then as Gale and Shapley proved in 1962:

**Proposition 1** *A stable matching of the population* $P$ *exists.*

[5] This difference in the way of resolving male and female indifference may seem arbitrary, but nothing important depends on it. Furthermore, although it is well known that lexicographic orderings can create problems for the representation of preferences by utility functions, at this stage we do not require utility functions, only preference orderings.

## 2.3 Uniqueness of Equilibrium

The preferences described above have a particular structure which ensures a unique stable matching. Consider two men, $m_i$ and $m_j$, and two women, $w_k$ and $w_l$ such that $m_i < m_j$ and $w_k < w_l$. Then it cannot be the case that $m_i$ prefers $w_l$ and $m_j$ prefers $w_k$; the shorter man cannot prefer the taller woman and at the same time the taller man prefer the shorter woman. All other combinations are possible but we can rule out the preferences shown in Figure 2. Similarly it cannot be the case

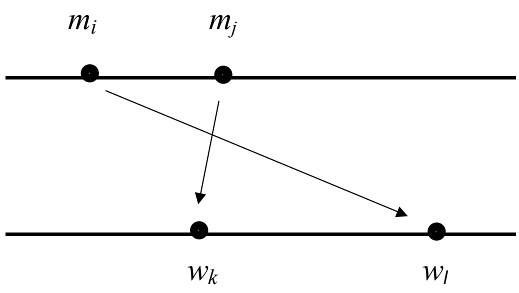


Figure 2: Preferences ruled out by the no crossing condition.

that $w_k$ prefers $m_j$ and $w_l$ prefers $m_i$. We now have:

**Proposition 2** *There exists a unique stable matching, $\mu^*$.*

To give an idea of how the unique stable matching may be found, we identify *fixed pairs* of the population $P$. A fixed pair is a couple $\{m, w\} \in M \times W$ who prefer each other over all other available partners. For any man $m$, consider his preferred woman in $W$, denoted $\theta(m)$. She is not necessarily the same height as $m$; this depends on the exact membership of the set $W$. Similarly, let $\gamma(w)$ denote $w$'s preferred man in $M$. Then $\gamma(\theta(m))$ is the preferred man of $m$'s preferred woman. If $m = \gamma(\theta(m))$, then since $m$ prefers $\theta(m)$ and $\theta(m)$ prefers $m$, the couple $\{m, \theta(m)\}$ are a fixed pair and must be matched in equilibrium.[6] The population $P$ must have at least one fixed pair, because the function $\gamma(\theta(m))$ is non-decreasing in $m$ (this follows from the No-Crossing Condition) and therefore has a fixed point in $M$ (which is finite). If we take out the fixed pairs of $P$, we are left with a sub-population $P' \subset P$, which must also satisfy the No Crossing Condition. Hence $P'$

[6] If $m \neq \gamma(\theta(m)$ think of $\gamma(\theta(m)$ as $m$'s *rival.* If $\theta(m) = w$, $m$ has no rival but as $m = \gamma(w)$, then $w = \theta(\gamma(m))$, so $w$ has no rivals either.

has at least one fixed pair, and these must also be matched in any stable matching of $P$, leaving a sub-population $P'' \subset P'$. We carry on in this way, until the whole population is matched.

As an example, suppose $M = \{1.41, 1.42, 1.48, 1.54, 1.77\}$, and $W = \{1.05, 1.12, 1.39, 1.75, 1.82\}$. Then the fixed pairs of $P$ are $\{1.41, 1.39\}$ and $\{1.77, 1.75\}$, the fixed pairs of $P'$ are $\{1.42, 1.12\}$ and $\{1.54, 1.82\}$, leaving $P'' = \{1.48, 1.05\}$, who must be paired with each other. The resulting matching is shown in Figure 3.

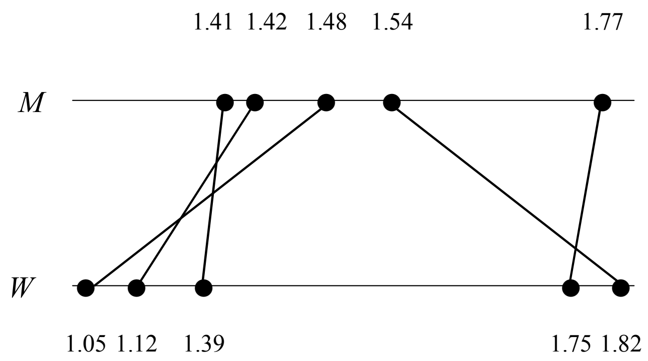


Figure 3: A unique equilibrium matching

## 2.4 Positive assortment

If individuals prefer similar partners, then it might be expected that in equilibrium they will tend to be matched with partners who are close to their ideal, and hence that taller men will be matched with taller women. But as Figure 3 shows, this intuition turns out to be misleading: "like attracts like" does not imply that "like is matched with like", and consequently need not result in positive assortment.

One measure of the degree of sorting is Spearman's rank correlation statistic, denoted by $\Lambda$. This captures the monotonocity of a matching. In the example of Figure 3, $\Lambda = 0.5$. Does the assumption of "like attracts like" put any restrictions on what values can be taken by $\Lambda$? Obviously, if $M = W$, then $\Lambda = 1$. In this case each man can find an identical and hence ideal woman, and vice versa, and the population $P$ forms immediately into $n$ perfectly matched fixed pairs. In contrast, $\Lambda = -1$ if either (i) the tallest woman is no taller than the shortest man, or (ii) the tallest man is no taller than the shortest woman. In the first instance, the population $P$

has only one fixed pair, the shortest man and the tallest woman, so they are matched in equilibrium; $P'$ has only one fixed pair, the second shortest man and the second tallest woman, so they are matched in equilibrium; and so on, until the tallest man is matched with the shortest woman. An example of this was given in Case 2 of the Introduction.

Thus no general restrictions on the degree of assortment are imposed by the assumption that "like attracts like". However, suppose that although women are on average shorter than men the tallest woman is taller than the shortest man; i.e. let $w_1 < m_1 < w_n < m_n$. There is then some "overlap" in height between the two sexes. If $n = 2$, we may still find that $\Lambda = -1$; for example if $M = \{1, 4\}$ and $W = \{3, 6\}$, there is one fixed pair of $P$, $\{4, 3\}$, leaving $\{1, 6\}$ as the other matched pair. But in a large population, there may be a significant number of men and women between $m_1$ and $w_n$, and so we might expect to find some fixed pairs in this range. If $P$ has at least two fixed pairs, then $\Lambda > -1$. This follows from the No Crossing Condition: within the sub-population who constitute the fixed pairs of $P$ there must be perfect positive assortment; otherwise we would be able to find preferences of the type illustrated in Figure 1. Thus the overall population $P$ cannot display perfect negative assortment.

Alternatively, suppose that men and women are on average the same height, but there is more variation in one sex; e.g. $m_1 < w_1 < w_n < m_n$. For $n = 2$, this population must sort positively: $m_1$ matches with $w_1$, and $m_2$ with $w_2$, so $\Lambda = 1$. For large $n$, there may be many men whose height is between $w_1$ and $w_n$, and who then find women with whom to form fixed pairs. Again this sub-population must sort positively. But we are now left with a sub-population $P'$ consisting of (i) men who are either shorter than $w_1$ or taller than $w_n$, or who did not find a woman in the range $w_1$ to $w_n$; and (ii) women between $w_1$ and $w_n$. It cannot be an equilibrium for $P'$ to sort positively, because that would imply that the shortest woman in $P'$ would match with $m_1$ and such a matching would be blocked by her and the tallest man shorter than $w_1$. If $P'$ has some negative sorting then the overall population $P$ cannot display perfect positive assortment and $\Lambda < 1$.

These arguments suggest two ways in which we can put some structure on the determination of the degree of assortment, and these are pursued in the next section. Firstly, we focus on the extent and nature of "overlap" between the sets $M$ and $W$. If there are intervals of $R$ which contain numbers of both men and women, this is where we will find the fixed pairs of $P$, even if the concentrations of men and women in any particular interval are different. This is important not only because within the fixed pairs of $P$ there must be positive assortment, but also because it is thorough the emergence of fixed pairs in successively smaller sub-populations that we can find the unique stable matching and hence analyse its characteristics. Secondly, we analyse large sets $M$ and $W$. This enables us to impose a degree of regularity on the distribution of height between the two sexes. Throughout we continue to assume that $n$ is finite, but of special interest is what happens as $n$ tends to infinity. We see that for a particular class of distributions we get remarkably simple and revealing results on the determination of the degree of assortment.

# 3 SORTING IN LARGE POPULATIONS

## 3.1 Fixed pairs in overlapping populations

We begin by making precise the idea that in intervals of $R$ containing large numbers of both men and women we will find a large number of fixed pairs. Consider an interval $Z = [x, y] \subset R$. Let $M_Z = Z \cap M$, the set of men in $Z$; similarly $W_Z = Z \cap W$. We assume that $M_Z$ and $W_Z$ both have more than two elements. Suppose that within $M_Z$ height is *evenly distributed* amongst men, in the sense that there is a constant difference $d_m$ in height between successively taller men; similarly, within $W_Z$ height is *evenly distributed* amongst women, with a constant difference in height $d_w$. If $d_m >$ (*resp.* $<$) $d_w$ we say that men (resp. women) are *sparse in* $Z$ and women (resp. men) are *abundant in* $Z$. Note that $M_Z$ and $W_Z$ may or may not include $x$ or $y$, and that the two sets need have no elements in common. We are not making any assumptions about how height is distributed outside $Z$. As an example of two even distributions, let $Z = [1.80, 2.00]$, $M_Z = \{1.81, 1.84, ..., 1.99\}$, and $W_Z = \{1.804, 1.824, ..., 1.984\}$, so $d_m = 0.03$, $d_w = 0.02$ and men are sparse in $Z$.

To state the main proposition on matching between two evenly distributed groups of men and women, it is useful to define sets that explicitly exclude the shortest and tallest elements of $M_Z$ and $W_Z$. Let $\overline{M}_Z = \{m \in M_Z | \min_{v \in M_Z} v < m < \max_{v \in M_Z} v\}$ and $\overline{W}_Z = \{w \in W_Z | \min_{v \in W_Z} v < w < \max_{v \in W_Z} v\}$.

**Proposition 3** *(a) If $d_m > d_w$ then (i) each man in $\overline{M}_Z$ is a member of a fixed pair of $P$, and is matched in equilibrium with a woman in $W_Z$; (ii) if a woman in $\overline{W}_Z$ is a member of a fixed pair of $P$, then she is matched in equilibrium with a man*

*in $M_Z$.*
*(b) If $d_m < d_w$ then (i) each woman in $\overline{W}_Z$ is a member of a fixed pair of $P$, and is matched in equilibrium with a man in $M_Z$ ; (ii) if a man in $\overline{M}_Z$ is a member of a fixed pair of $P$, then he is matched in equilibrium with a woman in $W_Z$*
*(c) If $d_m = d_w$ then (i) each man in $\overline{M}_Z$ is a member of a fixed pair of $P$ and is matched in equilibrium with a woman in $W_Z$; (ii) each woman in $\overline{W}_Z$ is a member of a fixed pair of $P$, and is matched in equilibrium with a man in $M_Z$.*

Within the interval $Z$, each member of the sparse sex, with the possible exception of the shortest and tallest, pairs with their preferred partner, as do some members of the abundant sex. But there are not enough of the sparse sex to go round, and so some of the abundant sex are left unsatisfied, in the sense that they do not match with their preferred partner. To illustrate this, consider the example above in which $Z = [1.80, 2.00]$, and $d_m = 0.03$ and $d_w = 0.02$. Then $\overline{M}_Z = \{1.84, 1.87, 1.90, 1.93, 1.96\}$, who are matched with women $\{1.844, 1.864, 1.904, 1.924, 1.964\}$ respectively, these two sets forming five fixed pairs of $P$. Note that there are women $\{1.824, 1.884, 1.944\}$, all in $\overline{W}_Z$, who are not members of fixed pairs of $P$. They make up a proportion $\frac{3}{8}$ of $\overline{W}_Z$, which is close to the proportion $1 - \frac{d_w}{d_m} = \frac{1}{3}$. Deviations from $1 - \frac{d_w}{d_m}$ arise since we are dealing with finite sets. To make this notion precise, consider what happens as we increase the membership of the sets $M_Z$ and $W_Z$, keeping the interval $Z = [x, y]$ fixed, and maintaining the assumption of two even distributions. We let $d_m$ and $d_w$ get smaller and smaller, but keep constant the ratio of $d_m$ to $d_w$. Denote by $\pi(S)$ the proportion of agents in set $S$ who are members of a fixed pair of $P$. Then

**Proposition 4** *Given the interval $Z = [x, y]$, as $d_m$ and $d_w$ both tend to zero with $d_w/d_m = \delta$,*
*(i) both $\pi(\overline{M}_Z)$ and $\pi(M_Z)$ tend to* $\min[1, \delta^{-1}]$;
*(ii) both $\pi(\overline{W}_Z)$ and $\pi(W_Z)$ tend to* $\min[1, \delta]$.

## 3.2 Sorting as correlation

Over the interval $Z = [x, y]$, all of the sparse sex (except possibly the shortest and tallest) match with their preferred partner. Taller men in $\overline{M}_Z$ prefer and match with taller women, so within the couples forming these fixed pairs there is positive assortment. But for values of $d_m$ and $d_w$ close to zero, the partners in a fixed pair must be very close to each other. This has the consequence that if we take as a measure of sorting the correlation coefficient $r$ between partners' heights then, restricting ourselves to the fixed pairs of $P$ where both partners are in $Z$, this measure is close to 1. Formally, for any matching $\mu$ we can compute the correlation coefficient between the height of those individuals in any set $S = \{p_1, p_2, ..., p_{\#S}\}$ and the height of their partners $\mu(S) = \{\mu(p_1), \mu(p_2), ..., \mu(p_{\#S})\}$, where $S$ contains only men or only women:

$$r(S, \mu) = \frac{Cov(S, \mu(S)),}{[Var(S)Var(\mu(S))]^{0.5}}, \tag{1}$$

where

$$Cov(S, \mu(S) = \frac{1}{\#S} \sum_{p_i \in S} (p_i - \overline{p}_S)(\mu(p_i) - \overline{\mu}_S),$$

$$Var(S) = \frac{1}{\#S} \sum_{p_i \in S} (p_i - \overline{p}_S)^2, \; \overline{p}_S = \frac{1}{\#S} \sum_{p_i \in S} p_i$$

$$Var(\mu(S)) = \frac{1}{\#S} \sum_{p_i \in S} (\mu(p_i) - \overline{\mu}_S)^2 \text{ and } \overline{\mu}_S = \frac{1}{\#S} \sum_{p_i \in S} \mu(p_i)$$

If $S = M$, then $r(S, \mu)$ is the correlation coefficient between the height of all men and their partners; clearly $r(M, \mu) = r(W, \mu)$, which we denote more simply by $r(\mu)$. We also write $r(S, \mu^*)$ as $r^*(S)$, and $r^*(M) = r(\mu^*)$ as $r^*$. We now have for the case where height is evenly distributed within $M_Z$ and $W_Z$:

**Proposition 5** *Given the interval $Z = [x, y]$, as $d_m$ and $d_w$ both tend to zero with $d_w/d_m = \delta$,*
*(i) if $\delta \leq 1$, then $r^*(\overline{M}_Z)$ tends to 1;*
*(ii) if $\delta \geq 1$, then $r^*(\overline{W}_Z)$ tends to 1.*

## 3.3 Sorting between two evenly distributed populations

We now apply Propositions 3 to 5 in a simple context. We assume that height is evenly distributed amongst both $M$ and $W$. It is useful to express each distribution in terms of its mean and spread. Thus

$$m_i = \bar{m} + s_m(\frac{i-1}{n-1} - \frac{1}{2})$$

where $\overline{m}$ is the average height of the men in $M$ and $s_m$ is the difference in height between the tallest and shortest man; thus the difference $d_m$ between succesively taller men is $s_m/(n-1)$. Similarly,

$$w_i = \bar{w} + s_w(\frac{i-1}{n-1} - \frac{1}{2})$$

so $d_w = s_w/(n-1)$. The assumption of two even distributions has the advantage that one obvious measure of the degree of assortment between the two sexes, Spearman's rank correlation statistic $\Lambda$, coincides with the correlation coefficient $r$.

The question now is: what determines the degree of assortment of the unique stable matching, as measured by $r^*$? The answer depends largely on the degree of overlap between the two distributions. There are three possible types of equilibria. In what follows I assume without loss of generality that $s_m \geq s_w$.

### 3.3.1 Type 1 matching: no overlap $w_n \leq m_1$ or $m_n \leq w_1$

If $w_n \leq m_1$ the tallest woman is no taller than the shortest man. In this case $m_i$ is the $i^{th}$ most preferred man of each woman, and $w_i$ is the $i^{th}$ least preferred woman of each man, so the sequence of fixed pairs that generates the equilibrium matching is $(m_1, w_n)$, $(m_2, w_{n-1})$, ..., $(m_{n-1}, w_2)$, $(m_n, w_1)$. As an example, let $n = 15$, $\overline{m} = 1.72$ and $s_m = 0.42$, so $M = \{1.51, 1.54, ..., 1.93\}$; and let $\overline{w} = 1.345$ and $s_w = 0.28$ so $W = \{1.205, 1.225, ..., 1.485\}$. This gives the matching pattern shown in Figure 4, from which it is clear that there is perfect negative assortment, with $r^* = -1$. Similarly, $r^* = -1$ if $m_n \leq w_1$.

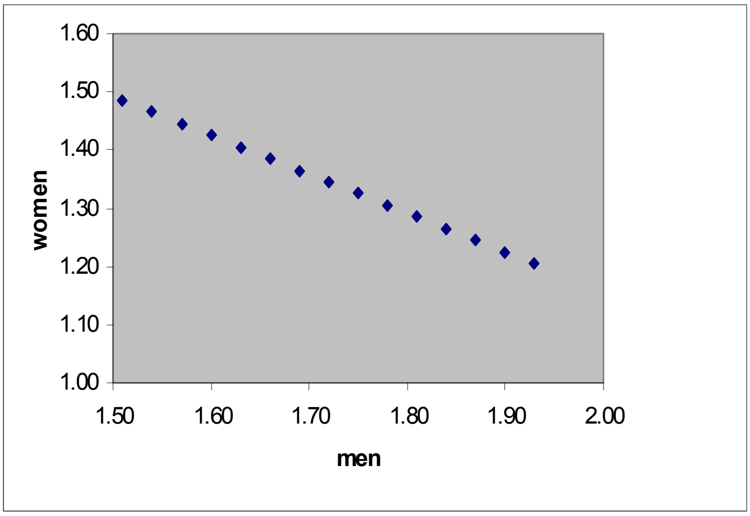


Figure 4: Type 1 matching

### 3.3.2 Type 2 matching: some overlap $w_1 < m_1 < w_n < m_n$ or $m_1 < w_1 < m_n < w_n$

We continue to denote by $Z$ the interval of overlap, so for a Type 2 matching $Z$ equals $[m_1,\, w_n]$ or $[w_1,\, m_n]$. We focus on the former case, where $w_1 < m_1 < w_n < m_n$. If $m_1 < w_1 < m_n < w_n$ and $Z = [w_1,\, m_n]$, the analysis follows a similar pattern. The number of men in $Z$ is denoted by $n_1$.

Proposition 3 shows that since $s_m \geq s_w$ each man in $\overline{M}_Z$ forms a fixed pair of $P$ with his preferred woman, who is in $W_Z$. Since $Z = [m_1,\, w_n]$ and the even distributions extend outside $Z$ to all men and women, we can establish the stronger result that each man in $M_Z$ forms a fixed pair with his preferred woman (although $\theta(m_1)$ may be shorter than $m_1$ and thus not in $W_Z$).[7] Within the fixed pairs of $P$ there can only be positive sorting, so as long as $n_1 \geq 2$ this overlap ensures a degree of positive assortment. In general there are more women than men in the range of overlap (because $s_m \geq s_w$), so the remainder of the population consists of women who range in height between $w_1$ and $w_n$, and men who range between $w_n$ and $m_n$. This subpopulation has no overlap between men and women, and therefore sorts negatively. Thus the equilibrium matching $\mu^*$ of the population $P$ has elements of positive and negative assortment. Overall, $-1 < r^* < 1$.

To illustrate this, we amend the example of a Type 1 matching above by increasing average female height, keeping all else the same; now $\overline{w} = 1.545$ so $W = \{1.405, 1.425, ..., 1.685\}$. The interval of overlap is now $Z = [1.51, 1.685]$; and $W_Z = \{1.525, 1.545..., 1.685\}$. The fixed pairs of this population are (1.51, 1.505), (1.54, 1.545), (1.57, 1.565), (1.60, 1.605), (1.63, 1.625), (1.66, 1.665), (1.69, 1.685). This leaves a subpopulation $P'$ of men $\{1.72, 1.75, ...1.93\}$ and women $\{1.405, 1.425, 1.445, 1.465, 1.485, 1.525, 1.585, 1.645, \}$. The tallest woman in $P'$ is shorter than the shortest man so they sort negatively. This gives the overall matching pattern shown in Figure 5, which displays predominantly negative assortment, with $r^* = -0.604$.

[7] We need to show that the least and greatest elements of $M_Z$ ($m_1$ and, say, $m_r$ respectively) have no rivals (as defined in Footnote 6). We follow the line of argument used to prove Proposistion 3. If $s_m > s_w$, then $\theta(m_1)$ is less than a distance $d_m/2$ from $m_1$ and more than $d_m/2$ from any other man; hence $m_1 = \gamma(\theta(m_1))$ and they form a fixed pair. This requires that there are no men shorter than $m_1$. Similarly, $m_r$ and $\theta(m_r)$ must be less than $d_m/2$ apart and hence form a fixed pair; this relies on the fact that if $w_n = \theta(m_q)$ then $w_n$ must be more than $d_m/2$ from $m_{q+1}$ (because all men, not just those in $Z$, are separated by a gap $d_m$). If $s_m = s_w$ and $d_m = d_w = d$, then either (A): $m_1$ is halfway between $w_t$ and $w_{t+1}$for some t, in which case $(m_1, w_t)$ is a fixed pair and, since $r = n - t$, so is $(m_q, w_{n-1})$; or (B) $m_1$ is less than $d_m/2$ from $\theta(m_1)$. in which case they form a fixed pair, as do $(\gamma(w_{n-1}), w_{n-1})$ and $(\gamma(w_n), w_n)$; either $\gamma(w_{n-1})$ or $\gamma(w_{n-1})$ must be the greatest element in $M_Z$.

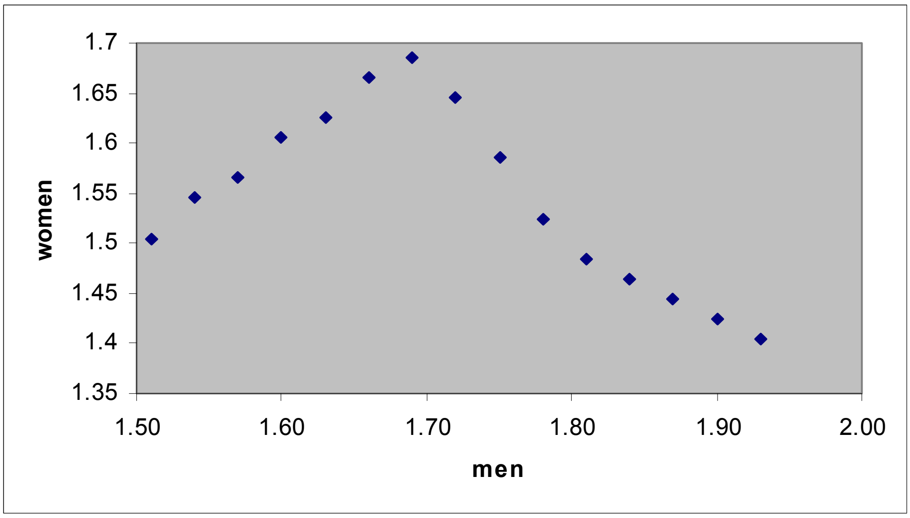


Figure 5: Type 2 matching

More generally, we can see that a type 2 matching $\mu^*$ has three parts: (i) $\mu_1^*$, the matching of all the $n_1$ men in the interval $Z = [m_1, w_n]$ with their preferred partners, each of whom, with the possible exception of $\theta(m_1)$, is in $Z$; (ii) $\mu_2^*$, the matching of the $n_2$ remaining women in $Z$ not matched with men in $Z$ by part (i) - these women are matched with the $n_2$ shortest men taller than $w_n$; and (iii) $\mu_3^*$, the matching of the remaining $n_3$ women, all of whom are shorter than $m_1$, with the $n_3$ tallest men. It is useful to label the sets of men and women matched by the component $\mu_i^*$ of $\mu^*$ as $M_i$ and $W_i$ respectively (thus $M_1 = M_Z$). These sets' location on the real line is illustrated in Figure 6, where the area of each rectangle is proportional to the size of the set that it represents.

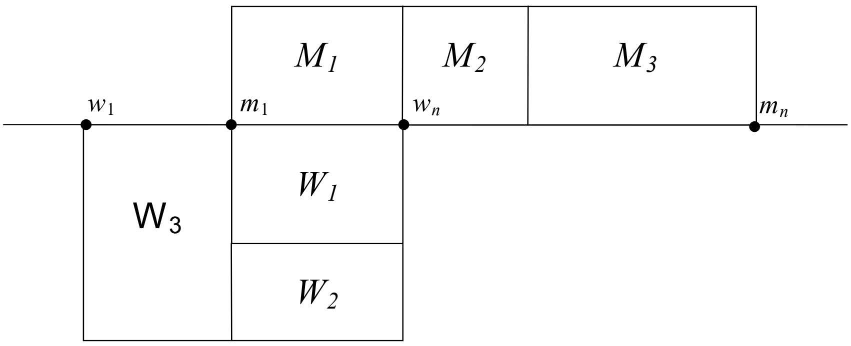


Figure 6: A population with a Type 2 matching; the men in $M_i$ match with the women in $W_i$, $i = 1, 2, 3$.

It is clear from Figure 5 that the correlation coefficient $r^*$ of a type 2 equilibrium matching with fixed values of the parameters $\overline{m}$, $\overline{w}$, $s_m$ and $s_w$ will in general be a cumbersome function of $n$. The following proposition allows a considerable simplification:

**Proposition 6** *For a type 2 matching with fixed values of the parameters $\overline{m}$, $\overline{w}$, $s_m$ and $s_w$ where $s_m \geq s_w$ and $\bar{w} - \frac{s_w}{2} < \bar{m} - \frac{s_m}{2} < \bar{w} + \frac{s_w}{2} w_n < \bar{m} + \frac{s_m}{2}$, as n tends to infinity*

$$r^* \to 2\left(\frac{s_m}{s_w}\right)^2 (\frac{s_m + s_w}{2s_m} - \frac{\bar{m} - \bar{w}}{s_m})^3 - 1. \tag{2}$$

*If $s_m \geq s_w$ and $\bar{m} - \frac{s_m}{2} < \bar{w} - \frac{s_w}{2} < \bar{m} + \frac{s_m}{2} < \bar{w} + \frac{s_w}{2} w_n$ then eq. (2) holds with $\bar{m} - \bar{w}$ replaced by $\bar{w} - \bar{m}$.*

#### 3.3.3 Type 3 matching: $m_1 \leq w_1 < w_n \leq m_n$

Here the distribution of women is contained within that of the men. From Proposition 3, each man in $\overline{M}_Z$ forms a fixed pair of $P$ with his preferred woman (as $s_m \geq s_w$); however, since $Z = [w_1, w_n]$ and the even distribution of $M$ extends to men shorter and taller than those in $Z$ , we can show that each man in $M_Z$ forms a fixed pair with his preferred woman[8]. This leaves a subpopulation consisting of (i) men who are either strictly shorter than $w_1$ or strictly taller than $w_n$, and (ii) women who are between $w_1$ and $w_n$ but were not matched with the men in this range. The unique stable matching for this subpopulation can be broken down into two parts, each with negative assortment. The argument is as follows.

Adopting notation similar to that used in the previous section, let $M_1$ and $W_1$ denote the sets whose members make up the fixed pairs of $P$, so $M_1 = M_Z$, and let $\mu_1^*$ denote the stable matching of the population $P_1 = M_1 \cup W_1$; thus $\mu_1^*(m) = \mu^*(m)$ for $m \in M_1$. Let $M_2$ be the set of men strictly shorter than $w_1$ and $M_3$ the set of men strictly taller than $w_n$, with $\#(M_2) = n_2$, and $\#(M_3) = n_3$; let $M_0 = M_2 \cup M_3$. Consider $W_0$, the set of women not matched with men in the interval $Z$; thus $\#(W_0) = n_2 + n_3$. Let $W_2$ be the set of the $n_2$ shortest women in $W_0$, and $W_3$ the set of the $n_3$ tallest women in $W_0$. These sets' location on the real

[8] Let $m_r$ and $m_t$ be the least and greatest elements of $M_Z$ respectively; then $w_1 \leq m_r < m_t \leq w_n$. If $s_m = s_w$ then $r = 1, t = n$, $w_1 = m_1$ and $w_n = m_n$, in which case $w_i$ forms a fixed pair with $m_i$ for all $i$. If $s_m > s_w$, $\theta(m_r)$ and $\theta(m_t)$ must be less than $d_m/2$ from $m_r$ and $m_t$ respectively, and thus have no rivals.

line is illustrated below, where I have assumed, without loss of any generality, that $n_2 \leq n_3$.

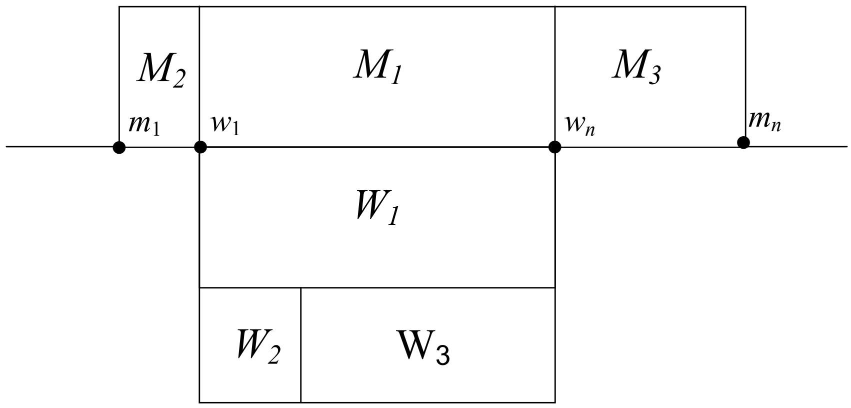


Figure 7: A population with a Type 3 matching; the men in $M_i$ match with the women in $W_i$, $i = 1, 2, 3$.

Consider the populations $P_2 = M_2 \cup W_2$ and $P_3 = M_3 \cup W_3$. Since $M_2$ and $W_2$ do not overlap, then $\mu_2^*$, the unique stable matching of $P_2$, displays perfect negative assortment. Similarly, $\mu_3^*$, the unique stable matching of $P_3$, also displays perfect negative assortment. Denote by $\mu'$ the unique stable matching of $P' = P_2 \cup P_3$. Then

**Lemma 1** *If $m \in M_2$ then $\mu'(m) = \mu_2^*(m)$; if $m \in M_3$ then $\mu'(m) = \mu_3^*(m)$.*

Thus the overall matching $\mu^*$ contains three parts such that $\mu_i^*(m) = \mu^*(m)$ for $m \in M_i$, $i = 1, 2, 3$. Where there is overlap, $\mu_1^*$ matches the fixed pairs of $P$ and therefore has positive assortment; the two others, $\mu_2^*$ and $\mu_3^*$, have negative assortment. In addition, there is a positive correlation between the means of the three groups of men and those of the three groups of women.

To illustrate a type 3 matching, we amend the example of a Type 2 matching and further increase the mean of $W$, keeping all else the same; now $\overline{w} = 1.705$ so $W = \{1.565, 1.585, ..., 1.845\}$. Then the fixed pairs in the area of overlap are {1.57, 1.565}, {1.60, 1.605}, {1.63, 1.625}, {1.66, 1.665}, {1.69, 1.685}, {1.72, 1.725}, {1.75, 1.745}, {1.78, 1.785}, {1.81, 1.805}, and {1.84, 1.845} leaving subsets of men $M_2 = \{1.51, 1.54\}$, who match negatively with $W_2 = \{1.585, 1.645\}$, and $M_3 = \{1.87, 1.90, 1.93\}$ who match negatively with $W_3 = \{1.705, 1.765, 1.825\}$. This gives the overall matching pattern shown in Figure 8, which displays predominantly positive assortment, with $r^* = 0.821$.

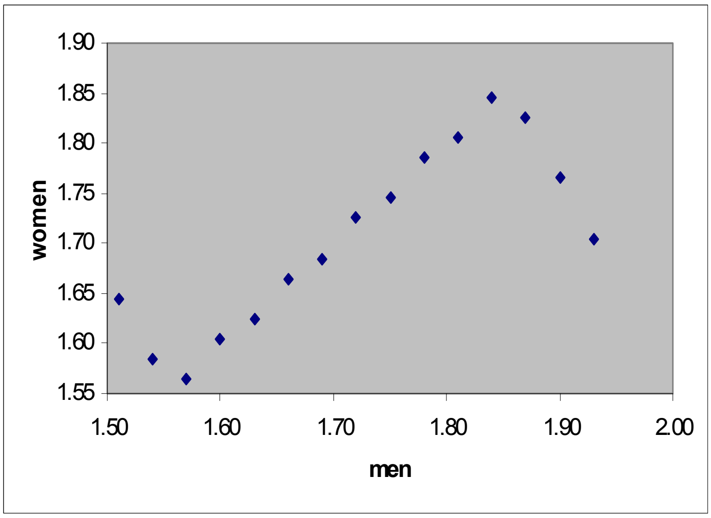


Figure 8: Type 3 matching

Note that in contrast to type 1 and type 2 matchings, a type 3 matching divides the population into two groups, shorter and taller, such that shorter men match only with shorter women, and taller men only with taller men, the definition of shorter being "less than height $h$", where $\max\{w|w \in W_2\} \leq h \leq \min\{w|w \in W_3\}$. As $n \rightarrow \infty$, $h$ tends to $\bar{w} - \frac{s_w}{s_m - s_w}(\bar{m} - \bar{w})$, and the proportion of shorter people to $\frac{1}{2} - \frac{s_m}{s_m - s_w}(\bar{m} - \bar{w})$. In Section 4, we interpret this division as a *stratification* of the population.

Akin to Proposition 6 we have:

**Proposition 7** *For $s_m \geq s_w$ and $\bar{m} - \frac{s_m}{2} < \bar{w} - \frac{s_w}{2} < \bar{w} + \frac{s_w}{2} w_n < \bar{m} + \frac{s_m}{2}$, as $n \rightarrow \infty$*

$$r^* \rightarrow \frac{s_m + s_w}{2 s_m} - \frac{6\left(\bar{m} - \bar{w}\right)^2}{s_m\left(s_m - s_w\right)} \tag{3}$$

#### 3.3.4 Determinants of the degree of assortment

To summarise the discussion so far, define $\sigma = \frac{s_w}{s_m}$ and $\Delta = \frac{\bar{m} - \bar{w}}{s_m}$. Then for $\sigma \leq 1$, as $n \rightarrow \infty$, $r^* \rightarrow \rho = f(\sigma, \Delta)$ where

$$f(\sigma, \Delta) = \begin{cases} -1 & \text{if } \Delta \leq -\frac{1+\sigma}{2} & \text{(Type 1)} \\ \frac{2}{\sigma^2}\left(\frac{1+\sigma}{2} + \Delta\right)^3 - 1 & \text{if } -\frac{1+\sigma}{2} < \Delta < -\frac{1-\sigma}{2} & \text{(Type 2)} \\ \frac{1+\sigma}{2} - \frac{6}{1-\sigma}\Delta^2 & \text{if } -\frac{1-\sigma}{2} \leq \Delta < \frac{1-\sigma}{2} & \text{(Type 3)} \\ \frac{2}{\sigma^2}\left(\frac{1+\sigma}{2} - \Delta\right)^3 - 1 & \text{if } \frac{1-\sigma}{2} \leq \Delta < \frac{1+\sigma}{2} & \text{(Type 2)} \\ -1 & \text{if } \frac{1+\sigma}{2} \leq \Delta & \text{(Type 1)} \end{cases} \tag{4}$$

If women vary more in height than men, then equation (4) holds with $\sigma$ and $\Delta$ redefined as $s_m/s_w$ and $(\bar{w} - \bar{m})/s_w$ respectively. Note that $\rho$ is (i) a continuous function of $\sigma$ and $\Delta$; (ii) symmetric around $\Delta = 0$; (iii) increasing as $|\Delta|$ decreases. $\rho$ does not necessarily increase with $\sigma$, although it does for $\sigma$ close to 1, and equals 1 only if $\sigma = 1$ and $\Delta = 0$. Figure 9 graphs the relationship between $\rho$ and $\Delta$ for $\sigma = 0.001$, 0.5, and 1.

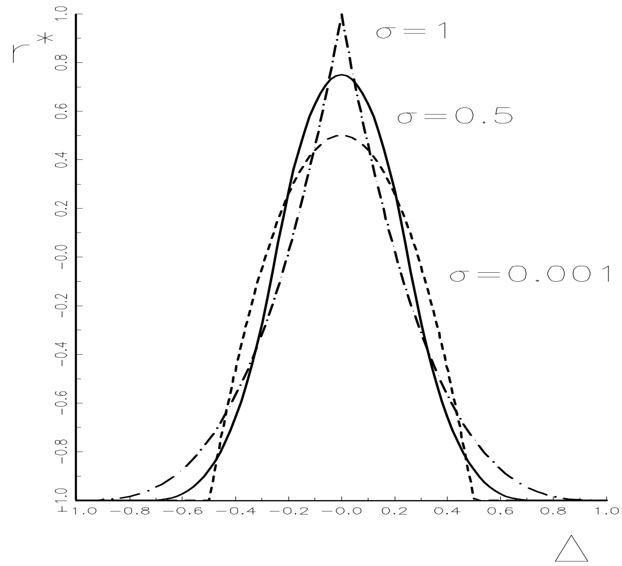


Figure 9: The degree of assortment, measured by the limiting value of the correlation coefficient as $n \to \infty$.

# 4 MATCHING AND WELFARE

## 4.1 Welfare and Correlation

The equilibrium matching $\mu^*$ cannot be Pareto dominated.[9] On efficiency gounds, then, there is no reason to object to an equilibrium that displays negative sorting, even if $r^* = -1$. But if matchings are evaluated using a welfare function which is neutral or averse to inequalities in individual utilities, then negative sorting will result in a low measure of social welfare if marginal disutility increases with the

[9]To see this, suppose $m_i$ and $w_j$ are not matched by $\mu^*$ but are matched by some alternative matching $\mu$ (if $\mu^* \neq \mu$, there must exist some such couple). Since $\mu^*$ cannot be blocked, it cannot be the case that $m_i$ prefers $w_j$ to $\mu^*(m_i)$ and $w_j$ prefers $m_i$ to $\mu^*(w_j)$. As preferences are strict, either $m_i$ or $w_j$ (or both) is worse off when matched by $\mu$.

difference $|x - y|$ between a spouse $x$ and his/her partner $y$ . To fix ideas, let $u_i$ be the utility of person $i$ and let

$$\Omega(\mu) = \frac{1}{2n} \sum_{i \in P} u_i$$

be a social welfare function which takes a simple average of individual utilities, which in turn depend on how the population $P$ is matched. If $m$ and $w$ are paired, we denote their combined utility, $u_m + u_w$, by $\phi(m, w)$. Then

$$\Omega(\mu) = \frac{1}{2n} \sum_{m \in M} \phi(m, \mu(m)).$$

We now apply standard results on supermodular functions: if the function $\phi(m, w)$ is supermodular (resp. submodular), then $\Omega(\mu)$ is maximised when matching displays perfect (resp. negative) positive assortment; see Topkis (1998), or Legros and Newman (2002).[10] If $\phi$ is twice differentiable, supermodularity (resp. submodularity) is equivalent to a non-negative (resp. non-positive) cross-partial derivative $\frac{\partial^2 \phi}{\partial m \partial w}$. Given $\Omega(\mu)$, therefore, the welfare consequences of matching depend on the form of individual utility functions. So far we have dealt only with individual preference orderings; for example a man $m$ ranks $w_i$ above $w_j$ if $|m - w_i| < |m - w_j|$, with any tie broken by a preference for the shorter woman. For given sets $M$ ands $W$, one way to represent such preferences is by the following utility functions: if $m \in M$ is matched with $w \in W$, his utility is

$$u_m(w) = -|w - m|^{\alpha} - \varepsilon w \tag{5}$$

and hers is

$$u_w(m) = -|m - w|^{\alpha} + \varepsilon m \tag{6}$$

where $\alpha > 0$. Here $\varepsilon$ is positive but sufficiently small that the male preference for shorter women is never strong enough to overturn the preference for a woman closer in height; i.e. for any $(m, w_i, w_j) \in M \times W^2$ such that $m - w_i > w_j - m > 0$

$$\varepsilon < \frac{(w_i - m_i)^{\alpha} - (m - w_j)^{\alpha}}{w_j - w_i}. \tag{7}$$

[10] These results are most commonly used in the case of transferable utility (TU). In the core of a TU game, total output/utility is maximised, so given such efficiency supermodularity implies positive sorting. In the framework that we consider, utility is non-transferable, so the direction of reasoning is different: given positive sorting, supermodularity implies welfare maximisation.

Similarly, in order for the female preference for taller men to act only as a tie-breaker we must also have

$$\varepsilon < \frac{(m_j - w)^\alpha - (w - m_i)^\alpha}{m_j - m_i} \tag{8}$$

for any $(m_i, m_j, w) \in M^2 \times W$ such that $m_j - w > w - m_i > 0$.[11]

Given these utility functions, $\phi(m, w) = \varepsilon(m - w) - 2(|m - w|)^\alpha$. We can therefore express social welfare as

$$\Omega(\mu) = \frac{\varepsilon}{2}(\bar{m} - \bar{w}) - \frac{1}{n}\sum_{m \in M}(|m - \mu(m)|)^\alpha. \tag{9}$$

$\phi(m,\, w)$ is supermodular if $\alpha \geq 1$, so in order to focus on the case where marginal disutility increases with the difference $|m - \mu(m)|$ we put $\alpha = 2$. This quadratic case allows a particularly crisp result on sorting and welfare. Rearranging equation (9) and bearing in mind the definition of $r(\mu)$, we have

$$\Omega(\mu) = a + br(\mu) \tag{10}$$

where

$$a = -Var(M) - Var(W) - (\bar{m} - \bar{w})^2 + \frac{\varepsilon}{2}(\bar{m} - \bar{w}) \tag{11}$$

$$b = 2\left(Var(M)Var(W)\right)^{0.5} \tag{12}$$

Both $a$ and $b$ are unaffected by the matching $\mu$, so we see that for a population with utility functions given by (5) and (6) with $\alpha = 2$, there is a positive linear relationship between social welfare and the degree of assortment as measured by $r(\mu)$.

Clearly the quadratic case is special; but even if utility functions are not given by (5) and (6) , positive sorting maximises an additive social welfare function such as $\Omega(\mu)$ if $\phi(m, w)$ is supermodular whatever the distribution of types, and maximises $r(\mu)$ whatever the form of the social welfare function or the distribution of types; negative sorting minimises such functions.[12]. This provides a broad justification for

[11] Implicitly this makes $\varepsilon$ depend on the actual sets $M$ and $W$ because for any positive $\varepsilon$, however small, one could construct a population with the utility functions above in which some man, for example, preferred the "wrong" (i.e. taller) woman despite a greater height difference compared to the "right" (i.e. shorter) woman. However for any finite sets $M$ and $W$ we can always choose $\varepsilon$ small enough to satisfy (7) and (8) .

[12] The correlation coefficient $r(\mu)$ can itself be expressed as $(c(\mu) - a)/b$, where $a$ and $b$ are given by equations (11) and (12), with $\varepsilon = 0$, and $c(\mu) = \frac{1}{2n}\sum_m (m - \mu(m))^2$. Since $(m - \mu(m))^2$ is a supermodular function of $m$ and $\mu(m)$, $c$ and hence $r$ are maximised (resp. minimised) by perfect positive (resp. negative) assortment.

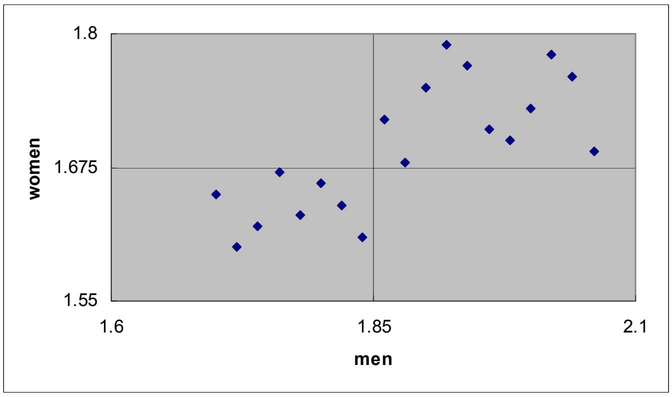


Figure 10: A stratified matching with two strata: men shorter than 1.85 match only with women shorter than 1.675.

taking $r(\mu)$ as a measure of welfare. However in what follows I explicitly assume an additive social welfare function and that equations (5) and (6) hold with $\alpha = 2$; thus $\Omega(\mu)$ increases if and only if $r(\mu)$ increases.

## 4.2 Welfare benefits of stratification

The results above imply that the equilibrium matching $\mu^*$ typically does not maximise social welfare. Indeed, in the case where $M$ and $W$ do not overlap, $\mu^*$ exhibits perfect negative sorting, yielding the lowest possible value of social welfare. We now consider the welfare consequences of a restriction on what matchings can form. In its most general formulation, this restriction operates as follows. Firstly, we partition each set $M$ and $W$ into $q$ non-empty subsets, $M_1$, $M_2$, $...M_q$ and $W_1$, $W_2$, $...W_q$ such that: (i) $\#M_i = \#W_i = n_i$, where $\sum_i n_i = n$ and (ii) if $m \in M_i$, $m' \in M_j$ and $i < j$, then $m < m'$; and if $w \in W_i$, $w' \in W_j$ and $i < j$, then $w < w'$. The subpopulation $P_i = M_i \cup W_i$ is called the $i^{th}$ *stratum* of $P$, and a partitioning of $M$ and $W$ in the way described above is a *stratification* of $P$. Secondly, we permit only matchings of $P$ that pair together individuals in the same stratum as each other. Such a matching is called a *stratified matching.* An example with $q = 2$, $n_1 = 8$ and $n_2 = 12$ is given in Figure 10.

It is possible that imposing a stratification may not be a binding constraint on the equilibrium matching. As Fgures 7 and 8 show, a Type 3 matching has two strata; men shorter than $h$ match only with women shorter than $h$, where $h$ divides $W_2$ from $W_3$. In this case, the population is endogenously stratified. But we might impose a different statification than that which emerges endogenously. Furthermore, Type 1 and 2 matchings are not stratified at all. What is the effect of imposing a stratification in these cases?

Our notion of equilibrium now applies to each stratum of $P$. The stratum $P_i$ is effectively a separate marriage market which satisfies the conditions for the existence of a stable matching. Moreover, $P_i$, being a subset of $P$, satisfies the No Crossing Condition and thus has a unique stable matching $\mu_i^*$. We denote by $r_i^*$ the correlation coefficient between the height of the men in $P_i$ and their partners' heights when they are matched by $\mu_i^*$. We denote by $\widetilde{\mu}$ the matching of $P$ implied by the $q$ matchings $\mu_1^*$, ..., $\mu_q^*$ (i.e. $m = \widetilde{\mu}(w)$ if $m = \mu_i^*(w)$ for some stratum $P_i$); similarly we denote by $\widetilde{r}$ the correlation coefficient between the height of the men in $P$ and their partners' heights when they are matched by $\widetilde{\mu}$ . Of course $\widetilde{\mu}$ and $\widetilde{r}$ depend on the stratification imposed on $P$.

To fix ideas, suppose all women are shorter than all men, so that $\mu^*$, the equilibrium matching when no stratification is imposed, exhibits perfect negative sorting. Then if $q \geq 2$, it must be the case that $\widetilde{r}$ is greater than $r^*$, the correlation coefficient associated with $\mu^*$. Although each of the matchings $\mu_1^*$, ..., $\mu_q^*$ considered separately exhibits perfect negative sorting, the stratification of $P$ induces some positive correlation: for $i < j$, all men in $M_j$ are taller than those in $M_i$, and their partners, who make up the set $W_j$, are taller than the women in $W_i$. If the partition of $P$ becomes finer (i.e. if we further stratify one or more of the strata of $P$) then clearly $\widetilde{r}$ increases further. The finest possible partition has $q = n$; in this case, and whether or not the sets $M$ and $W$ are overlapping, the $i^{th}$ tallest man matches with the $i^{th}$ tallest women, so sorting is perfectly positive.

In order to quantify the effects of stratification we return to the case where height is evenly distributed and $n$ tends to infinity. This allows us to take advantage of and extend the results already derived, and in particular to analyse the characteristics of the optimal stratification (which for a given $q$ maximises $\widetilde{r}$). We focus on two cases: (i) when $M$ and $W$ do not overlap, but $q$ varies; (ii) where $M$ and $W$ do overlap, and $q = 2$. We now establish some preliminary results.

However height is distributed and whatever the value of $n$, if $P$ is partitioned into $q$ strata then we can decompose the correlation coefficient $\widetilde{r}$ into $q + 1$ parts:

$$\widetilde{r} = \sum_{i=1}^{q} \omega_i r_i^* + \frac{Cov(\overline{m}_i, \overline{w}_i)}{[Var(M)Var(W)]^{0.5}},$$

where

$$\omega_i = \frac{n_i}{n}\frac{[Var(M_i)Var(W_i)]^{0.5}}{[Var(M)Var(W)]^{0.5}},$$

$$Cov(\overline{m}_i, \overline{w}_i) = \sum_{i=1}^{q} \frac{n_i}{n} (\overline{m}_i - \overline{m})(\overline{w}_i - \bar{w}),$$

$$\overline{m}_i = \frac{1}{n_i} \sum_{m_j \in M_i} m_j \text{ and } \overline{w}_i = \frac{1}{n_i} \sum_{w_j \in W_i} w_j.$$

We now let $n$ tend to infinity, with the relative size of each stratum tending to a given proportion of the population: for $i = 1, ...q$, as $n \to \infty$, we assume that $n_i/n \to \nu_i$, where $\sum_i^q \nu_i = 1$. Within each stratum, the matching $\mu_i^*$ will change as $n_i$ increases; but if height is evenly distributed within $M$ and $W$, with parameters $\overline{m}$, $\overline{w}$, $s_m$, and $s_w$, then as $n$ tends to infinity (i) $Var(M)$ and $Var(W)$ tend to $(s_m)^2/12$ and $(s_w)^2/12$ respectively; (ii) $Var(M_i)$ and $Var(W_i)$ tend to $(\nu_i s_m)^2/12$ and $(\nu_i s_w)^2/12$ respectively for all $i$; (iii) the correlation coefficient $r_i^*$ will tend to a limit $\rho_i = f(\sigma, \Delta_i)$ where $f(.,.)$ is given by equation (4), and $\Delta_i = Lim_{n\to\infty} (\overline{m}_i - \overline{w}_i)/(\nu_i s_m)$ $= \nu_i^{-1} \left\{\Delta + (1-\sigma)(\sum_j^{i-1} \nu_j - (1-\nu_i)/2)\right\}$. Thus as $n$ tends to infinity, $\widetilde{r}$ tends to

$$\widetilde{\rho} = \sum_{i=1}^{q} \nu_i^3 \rho_i + x, \tag{13}$$

where $x = 12\frac{Lim_{n\to\infty} Cov(\overline{m}_i, \overline{w}_i)}{s_m s_w}$. For any $n$, $Cov(\overline{m}_i, \overline{w}_i)$ depends on $n_1, n_2, , ...n_q$, but not on the matching $\widetilde{\mu}$ (which affects sorting within strata but not the strata means or their covariance). The value of $x$ is thus easily derived: if we consider the (typically non-equilibrium) matching that produces perfect positive sorting, then for any value of $n$ each $r_i$ equals 1, as well as the overall correlation coefficient $r$; i.e. $1 = \sum_{i=1}^q \omega_i + 12\frac{Cov(\overline{m}_i, \overline{w}_i)}{s_m s_w}$. Taking limits, we see that $x = 1 - \sum_{i=1}^q \nu_i^3$ and substituting into (13) we have:

$$\widetilde{\rho} = 1 - \sum_{i=1}^{q} \nu_i^3 (1 - \rho_i)$$

If the strata are of equal size, then $\nu_i = 1/q$ for all $i$ and $\widetilde{\rho} = 1 - q^{-2} + q^{-3} \sum_{i=1}^q \rho_i$. This is bounded below by $1 - 2q^{-2}$, which tends to 1 as $q$ tends to infinity. Thus the population can get arbitrarily close to the welfare maximum by a fine enough stratification.

#### 4.2.1 Optimal stratification when $M$ and $W$ do not overlap

If $M$ and $W$ do not overlap then neither do $M_i$ and $W_i$, so each $\mu_i$ is a type 1 matching. Thus as $n \to \infty$, $r_i$ tends to $-1$ for all $i$ and

$$\widetilde{\rho} = 1 - 2\sum_{i=1}^{q} \nu_i^3.$$

Finer stratification increases $\widetilde{\rho}$, since $\nu_i^3 + \nu_j^3 < (\nu_i + \upsilon_j)^3$ for $\nu_i$ and $\upsilon_j$ both strictly between 0 and 1. For fixed $q$, $\widetilde{\rho}$ is maximised when $\nu_i = 1/q$ for all $i$. In this case, $\widetilde{\rho} = 1 - 2/q^2$, the lower bound referred to above. For $q = 1$, $\widetilde{\rho}$ equals $-1$ and it approaches 1 only as $q$ increases without bound. Thus the welfare maximum is attainable only by an infinitely fine stratification of the population. However, note that having only two strata increases $\widetilde{\rho}$ from $-1$ to $\frac{1}{2}$.[13]

### 4.2.2 Optimal stratification when $q = 2$

Regardless of whether $M$ and $W$ overlap or not, when $q = 2$ and height is evenly distributed, then as $n$ tends to infinity $\widetilde{r}$ tends to

$$\widetilde{\rho} = 1 - {\nu_1}^3(1 - \rho_1) - {\nu_2}^3(1 - \rho_2). \tag{14}$$

If $M$ and $W$ overlap, then the values of $\rho_1$ and $\rho_2$ depend on the nature and extent of overlap and the relative size of the two strata. For example, it is possible that most but not all women are shorter than all men; then if $\nu_1$ is close to $\frac{1}{2}$, $M_1$ and $W_2$ overlap, but not $M_1$ and $W_1$, or $M_2$ and $W_2$. Then the unrestricted $\mu$ is a type 2 matching but $\widetilde{\mu}_1$ and $\widetilde{\mu}_2$ are both type 1 matchings with $\rho_1 = \rho_2 = -1$. But if $\nu_1$ is large enough, $M_1$ and $W_1$ do overlap and $\rho_1 > -1$. In general, $\rho_1$ and $\rho_2$ depend on the distribution parameters, $\overline{m}$, $\overline{w}$, $s_m$, and $s_w$, and the stratification, which for $q = 2$ can be represented by the value of $\nu_1$. More precisely, $\rho_i = f(\sigma, \Delta_i)$ where $\Delta_i = (\overline{m}_i - \overline{w}_i) / (\nu_1 s_m)$ and $f(., .)$ is given by equation (4). Note that

$$\Delta_1 = \frac{\Delta}{\nu_1} - \frac{(1-\sigma)(1-\nu_1)}{2\nu_1} \tag{15}$$

and

$$\Delta_2 = \frac{\Delta}{(1-\nu_1)} + \frac{(1-\sigma)\,\nu_1}{2\,(1-\nu_1)} \tag{16}$$

We can combine (4), (14), (15), and (**??**) to express $\widetilde{\rho}$ as a function $g(\nu_1; \sigma, \Delta)$. From this we can derive the optimal stratification correspondence $K(\sigma, \Delta) = \{\eta | \eta = \arg\max_{0 \le \nu_1 \le 1} g(\nu_1; \sigma, \Delta)\}$ and the maximised correlation function $h(\sigma, \Delta) = \max_{0 \le \nu_1 \le 1} g(\nu_1; \sigma, \Delta)$. The main properties of $g(\nu_1; \sigma, \Delta)$, $K(\sigma, \Delta)$ and $h(\sigma, \Delta)$ are set out in the three propositions below, with the formal analysis in Appendix 2.

**Proposition 8** *$g(\nu_1; \sigma, \Delta) \geq f(\sigma, \Delta)$ for all $0 < \nu_1 < 1$*

[13] This is reminiscent of McAfee's (2002) result that rationing electricity using two priority classes realises 75% of the value of using infiniteley many classes, the latter being first best but possibly administratively infeasible.

**Proposition 9** *(i) If $\Delta \neq 0$, $K(\sigma, \Delta)$ has a single element which is a continuous function $k(\sigma, \Delta)$ such that (a) if $0 < \sigma < 1$ and $0 < |\Delta| < 0.5$, then $k(\sigma, \Delta)$ is strictly decreasing in $|\Delta|$; (b) if $|\Delta| \geq 0.5$ or if $\sigma = 1$ then $k(\sigma, \Delta) = 0.5$; (c) $k(\sigma, \Delta) + k(\sigma, -\Delta) = 1$.*
*(ii) if $\Delta = 0$ and $0 < \sigma < 1$, then $K(\sigma, \Delta) = \left\{\widehat{k}(\sigma), 1 - \widehat{k}(\sigma)\right\}$, where $\widehat{k}(\sigma) =$ Lim $k(\sigma, \Delta)$ as $\Delta \to 0$ from above, $\widehat{k}(\sigma)$ is increasing in $\sigma$, $Lim_{\sigma \to 0}\widehat{k}(\sigma) = \frac{3}{4}$ and $Lim_{\sigma \to 1}\widehat{k}(\sigma) = \frac{5}{6}$;*
*(iii) if $\Delta = 0$, and $\sigma = 1$, $K(\sigma, \Delta) = \{v_1 | 0 \leq \nu_1 \leq 1\}$.*

**Proposition 10** *$h(\sigma, \Delta)$ is a continuous function such that (i) $h(\sigma, \Delta) = h(\sigma, -\Delta)$; (ii) $h(\sigma, \Delta) = 0.5$ if $|\Delta| > \frac{1}{2}$; (iii) $h(\sigma, \Delta) > 0.5$ if $|\Delta| < \frac{1}{2}$;(iii) $h(\sigma, \Delta) = 1$ if and only if $\Delta = 0$ and $\sigma = 1$; (iv) for $\Delta = 0$, $h(\sigma, \Delta)$ is increasing in $\sigma$;(v) for $\sigma < 1$, $h(\sigma, \Delta)$is increasing in $|\Delta|$ in a neighbourhood of $\Delta = 0$;*

Proposition 8 states that some stratification is never welfare reducing, even if the strata are of very different sizes. To understand how $K(\sigma, \Delta)$, the optimal stratification, and $h(\sigma, \Delta)$, the consequent correlation coefficient, depend on the underlying paramenters take two distributions that initially do not overlap because $\Delta \geq \frac{1+\sigma}{2}$, and consider the effects of reducing $\Delta$ keeping $\sigma$ constant at some value less than 1. With no overlap between $M$ and $W$, and with $q = 2$, $\widetilde{\rho}$ is maximised when $\nu_1 = \nu_2 = 0.5$. If $0.5 \leq \Delta < \frac{1+\sigma}{2}$, then as $\nu_1$ increases from 0 to 1, the matching types of $\mu_1^*$ and $\mu_2^*$ are (1, 2) respectively, then (1, 1) then (2, 1). The first and third of these would increase the correlation coefficient of one of the submatchings to a value more than $-1$, and have some effect on $x = 1 - {\nu_1}^3 - {\nu_2}^3$. By contrast, choosing $\nu_1 = \nu_2 = 0.5$ would mean that although $\mu_1^*$ and $\mu_2^*$ were both type 1 matchings, with correlation coefficients equal to $-1$, $x$ would attain its highest possible value of $\frac{3}{4}$. It is this effect that dominates, so $k(\sigma, \Delta) = h(\sigma, \Delta) = 0.5$. The optimal stratification when $\Delta = 0.5$ and $\sigma = 1$ is illustrated in Figure 11.[14]

As $\Delta$ falls below 0.5, choosing $\nu_1 = 0.5$ would mean that $\mu_1^*$ and $\mu_2^*$ were type 2 and type 1 matchings respectively, so $r_1^* > -1 = r_2^*$; it now becomes worthwhile to increase the size of the better matched stratum ($P_1$) and thus to reduce some of the more extreme mismatches in $P_2$; hence $k(\sigma, \Delta) > 0.5$. As $\Delta$ falls further, the benefit of having differently sized strata increases, even up to the point where $\Delta = 0$. In

[14]Because $\sigma = 1$, the unstratified matching $\mu^*$ has a discontinuity. In the region of overlap between $M$ and $W$ all men and women find an identical partmer, so the middle section apparent in Figure 5 has zero weight.

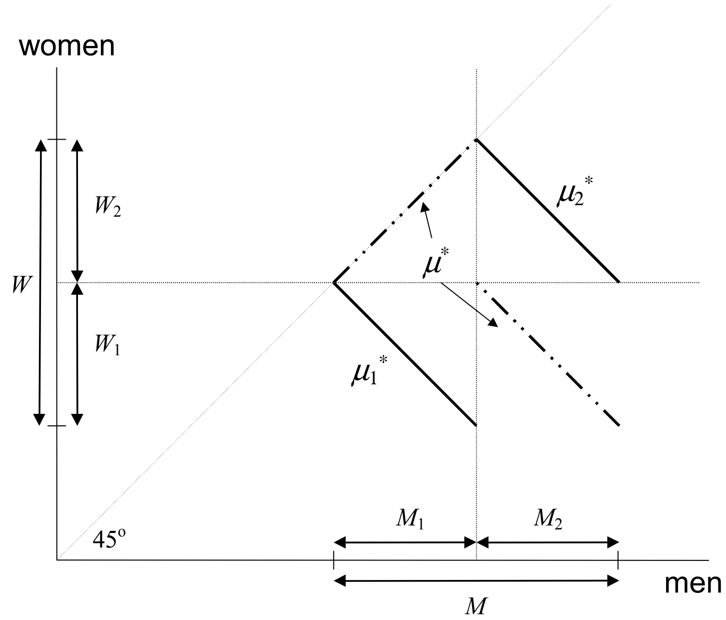


Figure 11: The optimal stratification when $\Delta = 0.5$ and $\sigma = 1$ has $\nu_1 = 0.5$

this case, the unrestricted matching $\mu^*$ has the shorter half of $M$ matching with the shorter half of $W$, so a symmetric stratification with $\nu_1 = 0.5$ would not increase welfare at all; any benefit from stratification can only be gained by having strata of different sizes. For $\Delta = 0$ it would not matter whether $P_1$ or $P_2$ were the larger; the function $g(\sigma, 0, \nu_1)$ is symmetric around $\nu_1 = 0.5$, achieving a minimum of $f(\sigma, 0)$ at $\nu_1 = 0$, 0.5, or 1, and a maximum of $h(\sigma, 0)$ at $\nu_1 = \widehat{k}(\sigma)$ or $\nu_2 = \widehat{k}(\sigma)$. If $\Delta$ is positive then the balance is in favour of having a larger $P_1$ than $P_2$. If $\Delta$ is negative, the balance is reversed; there is thus a discontinuity in the optimal stratification at $\Delta = 0$. As $\Delta$ approaches $-0.5$, so the optimal size of $P_2$ falls, with $1 - \nu_1^*$ getting closer to 0.5, and staying there for $\Delta < -0.5$. Figure 12 illustrates how the behaviour of how the optimal stratification varies with $\Delta$ for $\sigma = 0.001$, 0.5, and 0.99. It may be sumarised thus: if the sex that is the more varied in height is also on average the taller (e.g. $s_m > s_w$ and $\overline{m} > \overline{w}$), then there will be a relatively large stratum of short people ($v_1 > 0.5$) and a smaller stratum of tall people ($v_2 < 0.5$). [15]

Two special cases are worth mentioning. If $\Delta \neq 0$ then as $\sigma \to 1$, $k(\sigma, \Delta) \to 0.5$ and if $\sigma = 1$ then $k(\sigma, \Delta) = 0.5$; so the benefits of differently sized strata only exist if one sex is more varied in characteristics. Secondly, and in apparent contradiction to the first case, if $\Delta = 0$, then as $\sigma$ increases the optimal size of the larger stratum,

[15] For $\sigma = 1$ and $\Delta \neq 0$, $k(\sigma, \Delta) = 0.5$, and for $\sigma = 1$ and $\Delta = 1$ any stratification is optimal. Thus for $\sigma = 1$ the graph of $v_1{}^*$ against $\Delta$ would coincide with the two lines $v_1{}^* = 0$ and $\Delta = 0$.

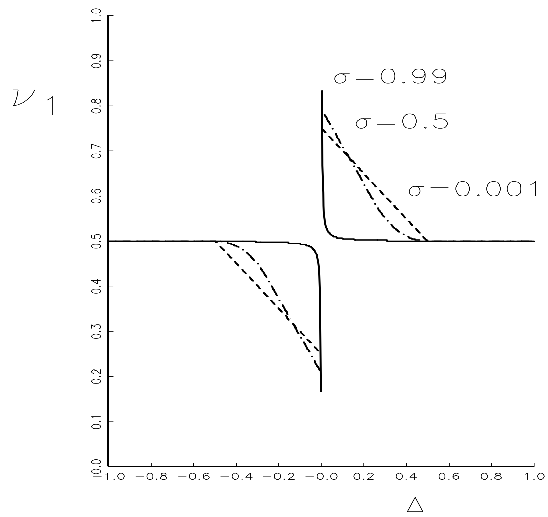


Figure 12: The optimal stratification as a function of $\Delta$.

$\widehat{k}(\sigma)$, increases with $\sigma$, and tends to $\frac{5}{6}$ as $\sigma$ approaches 1. But in this second case the benefits of any stratification are reduced as $\sigma$ approaches 1, since a correlation coefficient of $\frac{1+\sigma}{2}$ can be achieved by an unstratified matching. If $\Delta = 0$ and $\sigma = 1$, then $g(\nu_1; \sigma, \Delta) = 1$ for any value of $\nu_1$, including $\nu_1 = 0.5$.

Part (v) of Proposition 10 is perhaps somewhat counterintuitive. With an unstratified matching, welfare increases as $|\Delta|$ decreases, as 4 and Figure 9 show. By contrast, with an optimally chosen stratification and when $\Delta$ is close to zero, this is not the case unless $\sigma = 1$. Compared to $\Delta = 0$, a slight difference in the means of the two distributions actually reduces the cost of having differently sized strata, but at the margin it shifts the balance in favour of reducing the size of the larger stratum. However, for larger values of $|\Delta|$, further increases are welfare reducing. This is illustrated in Figure 13, which shows how the optimised correlation coefficient varies with $\Delta$ for $\sigma = 0.001$, 0.5, and 1.

### 4.2.3 Social Stratification

The results above show that a constraint on who can match with whom can have benefits in terms of welfare or aggregate payoff. We have taken utility functions with a particular form, but the essential assumption is that individuals experience an increasing marginal disutility from being badly matched. A welfare function with an aversion to inequality in individual utilities would reinforce the benefits of strat-

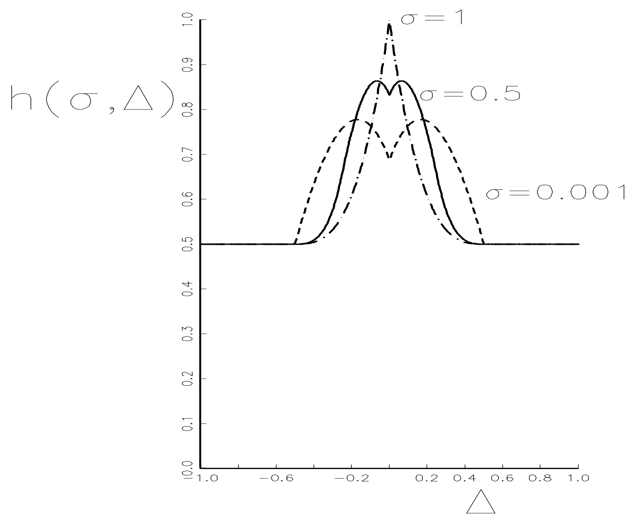


Figure 13: The correlation coefficient as a function $h(\sigma, \Delta)$ when the stratification is optimal.

ification. Since the unrestricted matching is Pareto efficient, a stratified matching secures a welfare improvement only by reducing the utility of some individuals. For example, in the case where all men are taller than all women, it forbids the matching of the tallest woman and the shortest man; instead, the tallest woman matches with the median man, and the shortest man with the median woman.

It is important to bear in mind that agents are horizontally not vertically differentiated (despite taking height as our main example of heterogeneity), and for the same partner taller individuals are not necessarily better off than shorter ones. It would therefore be quite wrong to interpret the strata as "upper class" and "lower class". Rather, it suggests the notion of strata as equally valued and possibly equally well-off communities, with matching only occurring within communities; we might call this "horizontal" stratification. It has important differences from the "vertical" stratification considered by some writers. For example, Durlauf (1996a) studies stratification when agents segregate into neighbourhoods according to income and education. Kremer and Maskin (1996) examine segregation of high and low skill workers into separate firms. Durlauf and Seshadri (2001) have a general model of coalition formation; in a stratified allocation, for any two coalitions all the members of one are more productive than all the members of the other. Economies with such differences almost inevitably have inequalities of income, and the focus of many of these studies is not inequality *per se* but how it is perpetuated by stratification, as in Benabou (1993) or (Durlauf, 1996a). This is in sharp contrast to the analysis presented here, which shows that stratification reduces inequalities in utility and increases welfare.

In the studies cited above vertical stratification arises endogenously, whereas in this paper, the mechanism by which constraints on matching arise is not explicitly modelled. The simplest justification for this is that stratification may arise as the result of technologically determined constraints on interaction due, for example, to transport costs or non-existent channels of communication. Then technical progress, by breaking down the barriers between communities, may reduce stratification and possibly bring about a fall in social welfare.

Alternatively, some problems of sorting arise in organisations where constraints on matching are explicit objects of choice. For example, corporate divisionalisation may make some working practices (e.g. pairings of workers) impossible and others easier.[16] The way in which a university is organised may encourage some forms of collaboration and discourage others. In some constitutional structures, such as devolution or federalism, the smaller political unit often has the power to impose constraints which amount to stratification; for example, the professional qualifications of teachers or lawyers may allow them to practise only in the region or state where they trained. By inhibiting mobility, such restrictions are sometimes regarded as welfare reducing. The analysis here suggests a case for the opposite.[17]

But even if stratification is not a technological fact of life or a policy variable, over time social pressures, norms, conventions, and sanctions might develop that would enforce a stratification. Indeed, such forces could be incorporated into agents' preferences, so that only "deviants" and not "normal people" would want to match with someone outside their own community. Taking this line of reasoning further, recall that the analysis shows that one of the reasons that stratification is welfare enhancing is that utility is not transferable. If utility is fully transferable and matched partners are faced with a utility possibility set such that $u_m + u_w = \phi(m, w)$, where $\phi$ is supermodular, then the population will sort positively and achieve maximum welfare. Legros and Newman (2003) have argued that economy-wide changes in the degree of transferability may help to explain changes in matching patterns.

[16] On corporate structure as a matching problem, see Legros and Newman (2003).

[17] More generally, we may think of various forms of *associational redistribution* discussed, for example, by Durlauf (1996b), as imposing a stratification to replace or improve that which arises endogenously.

If it is the case that the liberalisation of markets, deregulation, or the increased availability of credit, for example, have made utility more transferable, then there is less pressure for a stratification that constrains permissable partnerships, with a consequent weakening in the importance of community and social conventions.[18]

# 5 Concluding Remarks

Matching and sorting arise in many instances of social and economic activity. In some of these areas, it might be appropriate to assume that all participants on one side of the market agree on how to rank the participants on the other; in this case the matching equilibrium is unique and is characterised by positive sorting. But there are many situations where agents have different tastes, and where they would prefer to match with someone who has similar characteristics, i.e. where like attracts like. In this case the nature of the outcome is far from obvious. This paper show that when like attracts like there is a unique stable matching, but that there is no reason to expect it to display positive assortment. The degree of assortment depends on the distribution of characteristics among the two sides of the matching market. Where there is an overlap we will find couples who are well suited to reach other. But when such "fixed pairs" have been matched we may well be left with a sub-population where many agents prefer the same partner, who in turn may prefer someone else. By taking characteristics to be evenly, and in the limit uniformly, distributed this paper isolates the average mismatch between men and women, measured by $|\Delta|$, as the key determinant of assortment. The lower is $|\Delta|$, the greater is the degree of positive assortment.

By taking particular forms for the social welfare and utility functions, the paper shows that assortment - and hence welfare - can be increased by restrictions on matching. If matching occurs only within strata then in general the utility of some couples will be lower than it would otherwise be, but this is more than offset by the increase in the utility of others who would otherwise be in extreme mismatches. The analysis shows that if there are two strata, and they are chosen or have evolved optimally, then they will not necessarily be of equal size. If the sex that is taller on average is also the varied in height (e.g. for $\overline{m} > \overline{w}$ if $s_m > s_w$), then there will be

[18]But note that if agents are vertically not horizontally differentiated, then with non-transferable utility the population sorts positively, even with no constraints on matching. The assumption of "like attracts like" is central to the hypothesis that increased transferability reduces the need for stratification.

a relatively large stratum of short people and a smaller stratum of tall people.

This paper makes a number of important assumptions. These have delivered a tractable model and reasonably simple results, but the consequences of relaxing them remain to be explored. Two extensions of the model are worth a brief mention. Firstly, more general distributions of characteristics will produce a much wider variety of matching patterns than Types 1, 2, and 3. Nevertheless, the equilibrium matching can always be found by identifying the fixed pairs of sucessively smaller sub-populations. For the population $P$, it will still be in the regions of overlap between the two distributions that fixed pairs will be found. Since the fixed pairs of $P$ sort positively, the extent of overlap will be an important determinant of the degree of sorting. However, it seems likely that a simple characterisation of $\rho$ similar to equation (4) will only be available in special cases.[19]

Secondly, I have assumed that utility is absolutely non-transferable, so sorting is determined by the distribution of characteristics. At the other extreme of full transferability, sorting depends on the modularity of the function $\phi(m, w)$ giving the utilities of two matched individuals. This therefore leaves open the question of what happens to matching and sorting as the degree of transferability increases. For example, if $\phi$ is supermodular and the distributions of $m$ and $w$ do not overlap, then as we move from non- to full transferability, sorting will change from negative to positive. Investigating the properties of that transition is an interesting area for further research.

[19] Furthermore, it is only in the case of even or uniform distributions that the monotonicity of a matching, as measured by the the rank correlation $\Lambda$, coincides with the correlation coefficient $r$.

# APPENDIX 1

**Proof of Proposition 1.** Each person in $P$ has complete, reflexive, transitive and strict preferences over members of the opposite sex and would rather be married to anyone than remain single. The problem thus meets the conditions for existence set out in Gale and Shapley, (1962). ■

**Proof of Proposition 2.** The population $P$ satisfies the No Crossing Condition, as defined in Clark (2002), and so Theorem 2 in Clark (2002) applies. ■

**Proof of Proposition 3.** (a) (i) For each man $m \in \overline{M}_Z$, consider the open interval $Z_m = \{z|m - d_m/2 < z < m + d_m/2\}$. As $\overline{M}_Z$ excludes the shortest and tallest elements of $M_Z$, the shortest and tallest elements of $\overline{M}_Z$ must be at least a distance $d_m$ from $x$ and $y$ respectively, so $Z_m \subset Z$. Consider $\theta(m)$, $m$'s preferred woman in $W$. Since $d_m > d_w$, $Z_m$ must contain at least one element of $W$, hence $\theta(m) \in Z_m$.Furthermore, $Z_m$ is open and thus does not contain its boundary points, so $\gamma(w) = m$ for any woman $w \in Z_m$, since she is less than a distance $d_m/2$ from $m$ and more than $d_m$ from any other man. Hence $m = \gamma(\theta(m))$, so $m$ and $\theta(m) \in Z_m$ are a fixed pair and are therefore matched in equilibrium. Since $\theta(m) \in Z_m$ and $Z_m \subset Z$, then $\theta(m) \in Z$; but $\theta(m) \in W$, so $\theta(m) \in W \cap Z = W_Z$.(ii) Let $\{\widetilde{w}, \widetilde{m}\}$ be a pair such that $\widetilde{w} \in \overline{W}_Z$ and $\widetilde{m} \notin M_Z$. Then either $\widetilde{m} < x$, in which case $\widetilde{m}$ would prefer $\min_{v \in W_Z} v$ (the shortest element of $W_Z$) to $\widetilde{w}$, or $\widetilde{m} > y$, in which case $\widetilde{m}$ would prefer $\max_{v \in W_Z} v$ (the tallest element of $W_Z$) to $\widetilde{w}$. In either case, $\theta(\widetilde{m}) \neq \widetilde{w}$, so $\{\widetilde{w}, \widetilde{m}\}$ cannot be a fixed pair of $P$. Hence if $\widetilde{w}$ is a member of a fixed pair $\{\widetilde{w}, \widehat{m}\}$, then $\widehat{m} \in M_Z$.

(b) If women are sparse in $Z$, the proof of (a) applies, *mutatis mutandis.*

(c) Suppose now that $d_m = d_w = d$. We write $M_Z$ and $W_Z$ in order as $\{m_1^z, m_2^z, ..., m_q^z\}$ and $\{w_1^z, w_2^z, ..., w_r\}$ respectively, where $q = \#M_Z$ and $r = \#W_Z$. If $m_1^z = w_1^z$ then $q = r$ and $m_i^z$ equals and hence pairs with $w_i^z$ for $i = 1, 2, ...q$.To prove part (i) If $m_1^z \neq w_1^z$, then for any man in $\overline{M}_Z$ there exist men $\underline{m}$ and $\overline{m}$, both in $M_Z$, and women $\underline{w}$ and $\overline{w}$, both in $W_Z$, such that $\underline{m} < \underline{w} < m < \overline{w} < \overline{m}$. There are now two possibilities: (A) $m$ is equidistant between $\underline{w}$ and $\overline{w}$, $\underline{w}$ is equi-distant between $\underline{m}$ and $m$, and $\overline{w}$ is equi-distant between $m$ and $\overline{m}$. We have assumed that faced with two equidistant members of the opposite sex women prefer the taller man and men prefer the shorter woman, so $\{m, \underline{w}\}$ is a fixed pair. (B) If men and women are not equi-distant in the sense described in (A) above then $m$'s preferred woman in $P$, $\theta(m)$, who must be $\underline{w}$ or $\overline{w}$, is strictly less than $d/2$ in distance from $m$. Thus $m$ is also the nearest man to $\theta(m)$ so $m = \gamma(\theta(m))$ i.e.$\{m, \theta(m)\}$is a fixed pair.As for part (ii), for any woman in $\overline{W}_Z$ there exist women $\underline{w}$ and $\overline{w}$, both in $W_Z$, and men $\underline{m}$ and $\overline{m}$, both in $M_Z$, such that $\underline{w} < \underline{m} < w < \overline{m} < \overline{w}$. Then the same argument as for (i), *mutatis mutandis*, establishes that $\{\gamma(w), w\}$ is a fixed pair with $\gamma(w)$ equal either to $\underline{m}$ or $\overline{m}$. ■

**Proof of Proposition 4.** We provide a proof for the case where $\delta \geq 1$. The proof when $\delta \leq 1$ follows *mutatis mutandis* (i). By parts (i) and (iii) of Proposition 3 $\pi(\overline{M}_Z) = 1$ for any $d_m$, $d_w$ such that $d_m/d_w \geq 1$. By definition, $\#M_Z = \#\overline{M}_Z + 2$, (where $\#(S)$ denotes the number of elements in the set $S$), so $\pi(M_Z) \geq \frac{\#M_Z - 2}{\#M_Z}$. As $d_m \to 0$, $\#M_Z \to \infty$, so $\pi(M_Z) \to 1$. (ii) From Proposition 3, (a)(i) and(c)(i) each man in $\overline{M}_Z$ is in a fixed pair of $P$ with some woman in $W_Z$. Thus the number of women in $W_Z$ who are in fixed pairs is at least $\#(\overline{M}_Z)$ and the number of women in $\overline{W}_Z$ who are in fixed pairs is at least $\#(\overline{M}_Z) - 2$. From Proposition 3 (a)(iii) and (c)(iii) any woman in $\overline{W}_Z$ who is a member of a fixed pair of $P$ is matched in equilibrium with a man in $M_Z$, so the number of women in $\overline{W}_Z$ who are in fixed pairs can be no greater than $\#(M_Z)$. Dividing these lower and upper bounds by $\#(\overline{W}_Z)$, we have $\frac{\#(\overline{M}_Z)-2}{\#(\overline{W}_Z)} \leq \pi(\overline{W}_Z) \leq \frac{\#(M_Z)}{\#(\overline{W}_Z)}$. But for any values of $d_m$ and $d_w$, $(y-x)/d_m - 4 \leq \#(\overline{M}_Z) - 2 < \#(M_Z) \leq (y-x)/d_m + 1$ and $(y-x)/d_w - 4 \leq \#(\overline{W}_Z) \leq (y-x)/d_w - 1$ so $\leq \frac{(y-x)/d_m - 4}{(y-x)/d_w - 1} \leq \pi(\overline{W}_Z) \leq \frac{(y-x)/d_m + 1}{(y-x)/d_w - 4}$. As $d_m$ and $d_w$ tend to zero with $d_w/d_m = \delta$, these upper and lower bounds to $\pi(\overline{W}_Z)$ both tend to $\delta$, as does therefore $\pi(\overline{W}_Z)$. A similar argument establishes that $\frac{\#(\overline{M}_Z)}{\#(W_Z)} \leq \pi(W_Z) \leq \frac{\#(M_Z)+2}{\#(W_Z)}$, that $\frac{(y-x)/d_m - 2}{(y-x)/d_w + 1} \leq \pi(W_Z) \leq \frac{(y-x)/d_m + 3}{(y-x)/d_w - 2}$, and hence that $\pi(W_Z)$ tends to $\delta$. ■

**Proof of Proposition 5.** We assume without loss of generality that $\delta \leq 1$. The proof when $\delta \geq 1$ follows *mutatis mutandis.* If $\delta \leq 1$ and $\{m, w\}$ is a fixed pair such that $m \in \overline{M}_z$ then $|m - w| \leq d_m/2$. We exploit this to put lower and upper bounds on $r(\overline{M}_Z)$, both of which converge to 1. For any $d_m$, we can write $r(\overline{M}_Z)$ as

$$r(\overline{M}_Z) = \frac{\sum_{m_i \in S}(m_i(m_i \pm \frac{d_m}{2}) - \bar{m}_S(\bar{m}_S \pm \frac{d_m}{2})}{\left\{\sum_{m_i \in S}(m_i^2 - (\bar{m}_S)^2 \times \sum_{m_i \in S}[(m_i \pm \frac{d_m}{2})^2 - (\bar{m}_S \pm \frac{d_m}{2})^2]\right\}^{0.5}} \qquad (17)$$

where the notation $a \pm \frac{d_m}{2}$ means that $a$ lies between $a - \frac{d_m}{2}$ and $a + \frac{d_m}{2}$. Taking the highest and lowest values of $r(\overline{M}_Z)$ that satisfy eq.(17), and rearranging

$$\begin{aligned} & \frac{\sum_{m_i \in S}\{(m_i{}^2 - \bar{m}_S{}^2) - \bar{m}_S d_m\}}{\left\{\sum_{m_i \in S}(m_i{}^2 - \bar{m}_S{}^2) \times \{\sum_{m_i \in S}[(m_i{}^2 - \bar{m}_S{}^2) + 2m_i + \frac{d_m{}^2}{2}\}\right\}^{0.5}} \\ \leq \;& r(\overline{M}_Z) \\ \leq \;& \frac{\sum_{m_i \in S}\{(m_i{}^2 - \bar{m}_S{}^2) + \bar{m}_S d_m\}}{\left\{\sum_{m_i \in S}(m_i{}^2 - \bar{m}_S{}^2) \times \{\sum_{m_i \in S}[(m_i{}^2 - \bar{m}_S{}^2) - 2m_i - \frac{d_m{}^2}{2}\}\right\}^{0.5}} \end{aligned}$$

Dividing by $\#\overline{M}_Z$,

$$\begin{aligned} & \frac{Var(m, S) - \bar{m}_S d_m}{\{[Var(m, S)]^2 + [Var(m, S) \times (2\bar{m}_S d_m + \frac{d_m{}^2}{2})\}^{0.5}} \\ \leq \;& r(\overline{M}_Z) \\ \leq \;& \frac{Var(m, S) + \bar{m}_S d_m}{\{[Var(m, S)]^2 - [Var(m, S) \times (2\bar{m}_S d_m + \frac{d_m{}^2}{2})\}^{0.5}} \end{aligned}$$

As $d_m \to 0$, it is simple to show that $\overline{m}_S \to (x+y)/2$, and $Var(m, S) \to (x-y)^2/12$, (the variance of a uniform distribution on $[x, y]$ ) Thus $r(\overline{M}_Z)$ itself converges to 1. ■

**Proof of Lemma 1.** The statement of the Lemma clearly defines $\mu'$ as a matching of $P'$. Then we need to show that $\mu'$ cannot be blocked by any pair $(m, w) \in M_0 \times W_0$ such that $m \neq \mu'(w)$. Clearly $\mu'$ cannot be blocked by any pair $(m, w) \in M_2 \times W_2$, as this would imply that $\mu_2^*$ was not a stable matching of $P_2$; similarly, $\mu'$ cannot be blocked by any pair $(m, w) \in M_3 \times W_3$, as this would imply that $\mu_3^*$ was not a stable matching of $P_3$. Therefore if a pair $(m, w)$ can block $\mu'$, either (i) $(m, w) \in M_2 \times W_3$ or (ii). $(m, w) \in M_3 \times W_2$. If (i), then $m \in M_2$ prefers $w \in W_3$ to $\theta(m) \in W_2$; this is a contradiction, since all women in $W_2$ are nearer to all men in $M_2$ than all women in $W_3$. If (ii), then $m \in M_3$ prefers $w \in W_2$ to $\theta(m) \in W_3$; this is a contradiction, since all women in $W_3$ are nearer to all men in $M_3$ than all women in $W_2$. Thus $\mu'$ cannot be blocked and is therefore stable. ■

**Proof of Proposition 6.** As a preliminary, note that since $n_1/n$ is the proportion of men who are between $m_1 = \overline{m} - s_m/2$ and $w_n = \overline{w} + s_w/2$ in height and $n_3/n$ is the proportion of women who are between $w_1 = \overline{w} - s_w/2$ and $m_1 = \overline{m} - s_m/2$, then as $n \to \infty$,

$$\begin{bmatrix} n_1/n) \\ n_2/n) \\ n_3/n) \end{bmatrix} \to \begin{bmatrix} \frac{\overline{w}-\overline{m}+\frac{1}{2}(s_m+s_w))}{s_m} \\ \frac{(s_m-s_w)(\overline{w}-\overline{m}+\frac{1}{2}(s_m+s_w))}{s_m s_w} \\ \frac{\overline{m}-\overline{w}-\frac{1}{2}(s_m-s_w)}{s_w} \end{bmatrix}, \text{ denoted by } \begin{bmatrix} \upsilon_1 \\ \upsilon_2 \\ \upsilon_3 \end{bmatrix}.$$

**1. the limiting behaviour of** $\mu_1^*$. As $n \to \infty$, both $d_w$ and $d_m$ tend to 0. But $d_w/d_m = s_w/s_m$ for all $n$, so we may apply Propositions 4 and 5 with $\delta = s_w/s_m \leq 1$. In the limit, the men in $[m_1, w_n]$ are matched with a proportion $s_w/s_m$ of the women in $[m_1, w_n]$, with a correlation coefficient between men and their partners' heights of 1; i.e. in the limit $r(M_1) = 1$.

**2. the limiting behaviour of** $\mu_2^*$. For any $n$, the sets $M_2$ and $W_2$ match with perfect negative assortment, so the $\widetilde{n}th$ shortest man in $M_2$ matches with the $\widetilde{n}th$ tallest woman in $W_2$. In the limit $M_2$ ranges from $\bar{w} + s_w/2$ to $\bar{w} + s_w/2 + s_m\upsilon_2$; but for any $n$, male height within $M_2$ is distributed evenly, so in the limit for any fraction $\tilde{\pi}$ less than 1, the shortest $\tilde{\pi}$ of the men in $M_2$ range from $\bar{w} + s_w/2$ to $\bar{w} + s_w/2 + \pi s_m \upsilon_2$. From Proposition 4 $\pi(W_Z)$ (the limiting proportion of women in the interval $Z = [\bar{m} - s_m/2, \bar{w} + s_w/2]$ who are in a fixed pair of $P$ and thus in $W_1$) is $\delta = s_w/s_m$; this leaves a proportion $1 - s_w/s_m$ of the women in $Z$ who are in $W_2$. But Proposition 4 can be applied to any interval of the real line within which male and female height is evenly distributed with a given ratio $\delta = d_w/d_m$, so that that for *any* interval $Z' \subset [\bar{m} - s_m/2, \bar{w} + s_w/2]$, in the limit a proportion $1 - s_w/s_m$ of the women in $Z'$ are in $W_2$. Thus the height of women in $W_2$ is evenly distributed, but in general only in the limit. Since in the limit $W_2$ ranges from $\bar{m} - s_m/2$ to $\bar{w} + s_w/2$, this implies that the tallest $\tilde{\pi}$ of the women in $W_2$ range from

$\bar{w}+s_w/2-\tilde{\pi}(\bar{w}-\bar{m}+(s_w+s_m)/2)$ to $\bar{w}+s_w/2$. Thus a man of height $w_n+\pi s_m \upsilon_2$ is matched with a woman of height $\bar{w}+s_w/2-\tilde{\pi}(\bar{w}-\bar{m}+(s_w+s_m)/2)$. Recalling the expression for $\upsilon_2$ above, this means that the limiting behaviour of the sub-matching $\mu_2$ is characterised by a linear relationship $\mu^*(m)=a-bm$ where $b=\frac{s_m}{s_m-s_w}$, which implies a correlation coefficient between men and their partners' heights of -1; i.e. in the limit $r(M_2)=-1$.

**3. the limiting behaviour of** $\mu_3^*$.For any $n$, male height is distributed evenly within $M_3$ with a difference between succesively taller men of $d_m=s_m/(n-1)$; similarly female height is distributed evenly within $W_3$ with a difference between succesively taller women of $d_w=s_w/(n-1)$. These two sets sort perfectly negatively, so for any $n$ the behaviour of the sub-matching $\mu_3^*$ is characterised by a linear relationship $\mu^*(m)=a-bm$ where $b=\frac{s_m}{sw}$, which implies a correlation coefficient between men and their partners' heights of -1; i.e. $r(M_3)=-1$.

Now, for any $n$

$$r=\sum_{i=1}^{3}\omega_i r(M_i)+\frac{Cov(\overline{m}_i,\overline{w}_i)}{[Var(M)Var(W)]^{0.5}}, \tag{18}$$

where

$$\omega_i=\frac{n_i}{n}\frac{[Var(M_i)Var(W_i)]^{0.5}}{[Var(M)Var(W)]^{0.5}},$$

$$Cov(\overline{m}_i,\overline{w}_i)=\sum_{i=1}^{3}\frac{n_i}{n}\left(\overline{m}_i-\overline{m}\right)\left(\overline{w}_i-\bar{w}\right),$$

$$\overline{m}_i=\frac{1}{n_i}\sum_{m_j\in M_i}m_j \text{ and } \overline{w}_i=\frac{1}{n_i}\sum_{w_j\in W_i}w_j.$$

Within the sets $M$ and $W$ height is evenly distributed for any value of $n$. Thus as $n\to\infty$, $Var(M)\to\frac{s_m{}^2}{12}$ and $Var(W)\to\frac{s_w{}^2}{12}$.Within the sets $M_1$, $M_2$,$M_3$, and $W_3$ height is evenly distributed for any value of $n$. The discussion above of $\mu_2$ showed that the height of women in $W_2$ is evenly distributed, but in general only in the limit. A similar argument applies to $W_1$. Therefore as $n\to\infty$,

$$\begin{bmatrix}\overline{m}_1\\ \overline{m}_2\\ \overline{m}_3\\ \overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{bmatrix}\to\begin{bmatrix}\frac{2(\bar{m}+\bar{w})+s_w-s_m}{4}\\ \overline{w}+\frac{s_w+\upsilon_2 s_m}{2}\\ \bar{m}+\frac{s_m(1-\upsilon_2)}{2}\\ \frac{2(\bar{m}+\bar{w})+s_w-s_m}{4}\\ \frac{2(\bar{m}+\bar{w})+s_w-s_m}{4}\\ \frac{2(\bar{m}+\bar{w})-s_w-s_m}{4}\end{bmatrix} \text{ and } \begin{bmatrix}Var(M_1)\\ Var(M_2)\\ Var(M_3)\\ Var(W_1)\\ Var(W_2)\\ Var(W_3)\end{bmatrix}\to\begin{bmatrix}\frac{(\upsilon_1 s_m)^2}{12}\\ \frac{(\upsilon_2 s_m)^2}{12}\\ \frac{(\upsilon_3 s_m)^2}{12}\\ \frac{(\upsilon_1 s_m)^2}{12}\\ \frac{(\upsilon_1 s_m)^2}{12}\\ \frac{(\upsilon_3 s_w)^2}{12}\end{bmatrix}.$$

These limits all exist and are finite. Since the limits of $r(M_1)$, $r(M_2)$, and $r(M_3)$, are 1, $-1$, and $-1$ respectively, then as $n \to \infty$, $r \to r^*$ where

$$
\begin{aligned}
r^* \;=\; & \upsilon_1 \frac{(\upsilon_1 s_m)^2}{s_m s_w} - \upsilon_2 \frac{(\upsilon_2 s_m)(\upsilon_1 s_m)}{s_m s_w} - \upsilon_3 \frac{(\upsilon_3 s_m)(\upsilon_3 s_w)}{s_m s_w} \\
& + \{\upsilon_1 \frac{(2(\bar{m}+\bar{w}) + s_w - s_m)^2}{16} \\
& + \upsilon_2 \frac{(2\overline{w} + s_{w+}\upsilon_2 s_m)(2(\bar{m}+\bar{w}) + s_w - s_m)}{8} \\
& + \upsilon_3 \frac{(2\bar{m} + s_m(1-\upsilon_2))(2(\bar{m}+\bar{w}) - s_w - s_m)}{8} - \overline{m}\bar{w}\} / \frac{s_m s_w}{12}
\end{aligned}
$$

This boils down to

$$
r^* = 2\left(\frac{s_m}{s_w}\right)^2 \left(\frac{s_m + s_w}{2s_m} - \frac{\bar{m} - \bar{w}}{s_m}\right)^3 - 1 \tag{19}
$$

■

**Proof of Proposition 7.** Equation 18 still applies, but we must recompute its constituent parts, As a preliminary, note that since $n_1/n$ is the proportion of men who are between $w_1 = \overline{w} - s_w/2$ and $w_n = \overline{w} + s_w/2$ in height , $n_2/n$ is the proportion of women who are between $w_1 = \overline{w} - s_w/2$ and $m_1 = \overline{m} - s_m/2$, and $n_3/n$ is the proportion of women who are between $w_n = \overline{w} + s_w/2$ and $m_n = \overline{m} + s_m/2$, then as $n \to \infty$,

$$
\begin{bmatrix} n_1/n) \\ n_2/n) \\ n_3/n) \end{bmatrix} \to \begin{bmatrix} \frac{s_w}{s_m} \\ \frac{2(\overline{w}-\overline{m})+s_m-s_w}{2s_m} \\ \frac{2(\overline{m}-\overline{w})+s_m-s_w}{2s_m} \end{bmatrix}, \text{ which we denote by } \begin{bmatrix} \upsilon_1 \\ \upsilon_2 \\ \upsilon_3 \end{bmatrix}.
$$

**1. the limiting behaviour of** $\mu_1^*$.As $n \to \infty$, both $d_w$ and $d_m$ tend to 0. But $d_w/d_m = s_w/s_m$ for all $n$, so we may apply Propositionss 4 and 5 with $\delta = s_w/s_m \leq 1$. In the limit, the men in $[w_1, w_n]$ are matched with a proportion $s_w/s_m$ of the women in $[w_1, w_n]$, with a correlation coefficient between men and their partners' heights of 1; i.e. in the limit $r(M_1) = 1$.

**2. the limiting behaviour of** $\mu_2^*$.For any $n$, the sets $M_2$ and $W_2$ match with perfect negative assortment, so the $\widetilde{n}th$ tallest man in $M_2$ matches with the $\widetilde{n}th$ shortest woman in $W_2$. In the limit $M_2$ ranges from $\bar{m} - s_m/2$ to $\bar{w} - s_w/2$; but for any $n$, male height within $M_2$ is distributed evenly, so in the limit for any fraction $\tilde{\pi}$, the tallest $\tilde{\pi}$ of the men in $M_2$ range from $\bar{w} - s_w/2 - \tilde{\pi}(\bar{w} - \bar{m} + (s_m - s_w)/2)$ to $\bar{w} - s_w/2$. From Proposition 4 $\pi(W_Z)$, the limiting proportion of women in the interval $Z = [\bar{m} - s_m/2, \bar{m} + s_m/2]$ who are in a fixed pair of $P$ (and thus in $W_1$) is $\delta = s_w/s_m$; this leaves a proportion $1 - s_w/s_m$ of the women in $Z$ who are in $W_0$, and hence either $W_2$ or $W_3$. But Proposition 4 can be applied to any interval of the real line within which male and female height is evenly distributed with a given ratio $\delta = d_w/d_m$, so that that for *any* interval $Z' \subset Z$ in the limit proportions $s_w/s_m$ and $1 - s_w/s_m$ of the women in $Z'$ are in $W_1$ and $W_0$ respectively. Thus in the limit the height of women in $W_0$, and hence of those in $W_2$ and $W_3$, is evenly distributed. Since $W_0$ ranges from $\bar{w}-s_w/2$ to $\bar{w}+s_w/2$, this implies that in the limit $W_2$ ranges from $\bar{w}-s_w/2$ to $\bar{w}-s_w/2+\frac{v_2}{v_3}s_w$, the shortest $\tilde{\pi}$ of whom range from $\bar{w}-s_w/2$ to $\bar{w}-s_w/2+\tilde{\pi}\frac{v_2}{v_2+v_3}s_w$. A man of height $\bar{w}-s_w/2-\tilde{\pi}(\bar{w}-\bar{m}+(s_m-s_w)/2)$ is therefore matched with a woman of height $\bar{w}-s_w/2+\tilde{\pi}\frac{v_2}{v_2+v_3}s_w$. Recalling the expressions for $v_2$ and $v_2$ above, this means that the limiting behaviour of the sub-matching $\mu_2$ is characterised by a linear relationship $\mu(m)=a-bm$ where $b=\frac{s_w}{s_m-s_w}$, which implies a correlation coefficient between men and their partners' heights of -1; i.e. in the limit $r(M_2)=-1$.

**3. the limiting behaviour of** $\mu_3^*$.The argument here follows that regarding $\mu_2$ almost exactly. In the limit, a man of height $\bar{w}+s_w/2+\tilde{\pi}(\bar{m}-\bar{w}+(s_m-s_w)/2)$ is matched with a woman of height $\bar{w}+s_w/2-\tilde{\pi}\frac{v_3}{v_2+v_3}s_w$. Therefore $\mu_2$ is characterised by a linear relationship $\mu(m)=a-bm$ where $b=\frac{s_w}{s_m-s_w}$, implying a correlation coefficient between men and their partners' heights of -1; i.e. in the limit $r(M_2)=-1$.

Now, within the sets $M$ and $W$ height is evenly distributed for any value of $n$. Thus as $n\to\infty$, $Var(M)\to\frac{{s_m}^2}{12}$ and $Var(W)\to\frac{{s_w}^2}{12}$.Within the sets $M_1$, $M_2$, and $M_3$, height is evenly distributed for any value of $n$. The discussion above of $\mu_2$ showed that the height of women in $W_1$, $W_2$ and $W_3$ is evenly distributed, but in general only in the limit. Therefore as $n\to\infty$ ■

$$\begin{bmatrix}\overline{m}_1\\ \overline{m}_2\\ \overline{m}_3\\ \overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{bmatrix}\to\begin{bmatrix}\bar{w}\\ \frac{2(\bar{m}+\bar{w})-s_w-s_m}{4}\\ \frac{2(\bar{m}+\bar{w})+s_w+s_m}{4}\\ \bar{w}\\ \bar{w}-\frac{v_3s_ms_w}{2(s_m-s_w)}\\ \bar{w}+\frac{v_2s_ms_w}{2(s_m-s_w)}\end{bmatrix}\text{ and }\begin{bmatrix}Var(M_1)\\ Var(M_2)\\ Var(M_3)\\ Var(W_1)\\ Var(W_2)\\ Var(W_3)\end{bmatrix}\to\begin{bmatrix}{s_w}^2\\ (v_2s_m)^2\\ (v_3s_m)^2\\ {s_w}^2\\ (\frac{v_1v_2}{s_m-s_w})^2\\ (\frac{v_1v_3}{s_m-s_w})^2\end{bmatrix}\div 12.$$

These limits all exist and are finite. Since the limits of $r(M_1)$, $r(M_2)$, and $r(M_3)$ are 1, $-1$, and $-1$ respectively, then as $n\to\infty$, $r\to r^*$ where

$$\begin{aligned}r^* &= v_1\frac{{s_w}^2}{s_ms_w}-v_2\frac{(v_2s_m)(v_1v_2)}{s_ms_w(s_m-s_w)}-v_3\frac{(v_3s_m)(v_1v_3)}{s_ms_w(s_m-s_w)}\\ &\quad+\{v_1\bar{w}^2\\ &\quad+v_2\frac{(2(\bar{m}+\bar{w})(s_m-s_w)-{s_m}^2+{s_w}^2)\,(4\bar{w}(s_m-s_w)-2v_3s_ms_w)}{4(s_m-s_w)}\\ &\quad+v_3\frac{(2(\bar{m}+\bar{w})(s_m-s_w)+{s_m}^2-{s_w}^2)\,(4\bar{w}(s_m-s_w)+2v_2s_ms_w)}{4(s_m-s_w)}\\ &\quad-\bar{m}\bar{w}\}\div\frac{s_ms_w}{12}\end{aligned}$$

Again, this expression simplifies, and

$$r^*=\frac{s_m+s_w}{2s_m}-\frac{6\,(\bar{m}-\bar{w})^2}{s_m\,(s_m-s_w)}$$

# APPENDIX 2: properties of the optimal stratification

## 1. $\mu^*$ is a type 1 matching: $\frac{1+\sigma}{2} \leq |\Delta|$

In this case, whatever the value of $\nu_1$ neither $M_1$ and $W_1$, nor $M_2$ and $W_2$, overlap, so $\rho_1 = \rho_2 = -1$, $\widetilde{\rho} = 1 - 2{\nu_1}^3 - 2{\nu_2}^3$ and $\nu_1^* = \frac{1}{2}$. If the strata were of different sizes the larger stratum would involve some extreme mismatches, which would not be offset by the closer matches in the smaller stratum.

## 2. $\mu^*$ is a type 2 matching: $\frac{1-\sigma}{2} \leq |\Delta| < \frac{1+\sigma}{2}$

### 2.1 $\sigma \neq 1$

We focus on the case where $\Delta > 0$. The case where $\Delta < 0$ is symmetric. We consider two possibilities:

2.1.1 $\frac{1+\sigma^2}{2(1+\sigma)} \leq \Delta < \frac{1+\sigma}{2}$

Combining 4, 14, 15, and **??** we see that as $\nu_1$ increases from 0 to 1, the matching types of $\mu_1^*$ and $\mu_2^*$ are (1, 2) respectively, then (1, 1), then (2, 1). More precisely, $g(\nu_1; \sigma, \Delta) =$

$$\begin{cases} 6\nu_1 (1-\nu_1) - 1 + 2\sigma^{-2} \left(\frac{1+\sigma}{2} - v_1 - \Delta\right)^3 & \text{if } \ 0 \leq \nu_1 \leq \frac{1+\sigma}{2} - \Delta \\ 6\nu_1 (1-\nu_1) - 1 & \text{if } \ \frac{1+\sigma}{2} - \Delta < \nu_1 < \frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} \\ 6\nu_1 (1-\nu_1) - 1 + 2\sigma^{-2} \left(\sigma\nu_1 - \Delta + \frac{1-\sigma}{2}\right)^3 & \text{if } \ \frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} < \nu_1 \leq 1. \end{cases} \quad (20)$$

If $\nu_1 = 0$ or 1, then $g(\nu_1; \sigma, \Delta) = 2\sigma^{-2} \left(\frac{1+\sigma}{2} - \Delta\right)^3 - 1 = f(\sigma, \Delta)$. Note also that $g(\nu_1; \sigma, \Delta)$ is continuous in $\nu_1$ with a continuous first derivative $\partial g / \partial v_1$. For $0 \leq \nu_1 \leq \frac{1+\sigma}{2} - \Delta$, $\partial g / \partial v_1 > 0$ at both $\nu_1 = 0$ and $\nu_1 = \frac{1+\sigma}{2} - \Delta$. Since $g(\nu_1; \sigma, \Delta)$ is a cubic in $\nu_1$, with a negative coefficient on ${\nu_1}^3$ then $\partial g / \partial v_1 > 0$ for $0 \leq \nu_1 \leq \frac{1+\sigma}{2} - \Delta$.

Similarly, for $\frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} < \nu_1 \leq 1$, $g(\nu_1; \sigma, \Delta)$ is a cubic in $\nu_1$, with a positive coefficient on ${\nu_1}^3$. If $0.5 < \Delta$ then $\frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} > 0.5$ and $\partial g / \partial v_1 < 0$ at both $\nu_1 = \frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma}$ and $\nu_1 = 1$ so $\partial g / \partial v_1 < 0$ for $\frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} < \nu_1 \leq 1$. Thus $g(\nu_1; \sigma, \Delta)$ reaches a maximum in the interval $\frac{1+\sigma}{2} - \Delta < \nu_1 < \frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma}$. Hence, for $0.5 < \Delta < \frac{1+\sigma}{2}$, (i) $K(\sigma, \Delta) = \{0.5\}$; (ii) at this stratification the matching types of $\mu_1$ and $\mu_2$ are (1, 1) respectively; and (iii) $h(\sigma, \Delta) = 0.5$.

If $\frac{1+\sigma^2}{2(1+\sigma)} \leq \Delta \leq$ (*resp.* $<$) 0.5, then $\frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} \leq$ (*resp.* $<$) 0.5 and $\partial g / \partial v_1 \geq$ (*resp.* $>$) 0 at $v_1 = \frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma}$; but at $v_1 = 1$, $\partial g / \partial v_1 < 0$. Since, for $\frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} < \nu_1 \leq 1$, $g(\nu_1; \sigma, \Delta)$ is a cubic in $\nu_1$ with a positive coefficient on ${\nu_1}^3$, it reaches a local maximum in the interval $[\frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma}, 1]$ which (since $\partial g / \partial v_1 < 0$ for $0 < \nu_1 < \frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma}$) is a maximum over the interval [0, 1]. Thus for $\frac{1+\sigma^2}{2(1+\sigma)} \leq \Delta \leq 0.5$, the set $K(\sigma, \Delta)$ has a single element $k(\sigma, \Delta)$ defined by the first order condition $\partial g / \partial v_1 = 0$, with $k(\sigma, 0.5) = 0.5$ and $h(\sigma, 0.5) = 0.5$. Elementary calculus then shows that for $\frac{1+\sigma^2}{2(1+\sigma)} \leq \Delta < 0.5$ (i) $\partial k(\sigma, \Delta) / \partial \Delta$ exists and is strictly negative, implying $k(\sigma, \Delta) > 0.5$ so that at this stratification the matching types of $\mu_1$ and $\mu_2$ are (2, 1) respectively; (ii) $\partial h(\sigma, \Delta) / \partial \Delta$ exists and is strictly negative, implying $h(\sigma, \Delta) > 0.5$.

**2.1.2** $\frac{1-\sigma}{2} \leq \Delta \leq \frac{1+\sigma^2}{2(1+\sigma)}$ In this case, as $\nu_1$ increases from 0 to 1, the matching types of $\mu_1$ and $\mu_2$ are (1, 2) respectively, then (2, 2), then (2, 1). More precisely, $g(\nu_1; \sigma, \Delta) =$

$$\begin{cases} 6\nu_1(1-\nu_1) - 1 + 2\sigma^{-2}\left(\frac{1+\sigma}{2} - v_1 - \Delta\right)^3 & \text{if } 0 \leq \nu_1 \leq \frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} \\ 6\nu_1(1-\nu_1) - 1 + 2\sigma^{-2}\left(\frac{1+\sigma}{2} - v_1 - \Delta\right)^3 & \\ \quad + 2\sigma^{-2}\left(\sigma\nu_1 - \Delta + \frac{1-\sigma}{2}\right)^3 & \text{if } \frac{\Delta}{\sigma} - \frac{1-\sigma}{2\sigma} \leq \nu_1 \leq \frac{1+\sigma}{2} - \Delta \\ 6\nu_1(1-\nu_1) - 1 + 2\sigma^{-2}\left(\sigma\nu_1 - \Delta + \frac{1-\sigma}{2}\right)^3 & \text{if } \frac{1+\sigma}{2} - \Delta \leq \nu_1 \leq 1. \end{cases} \tag{21}$$

Again, if $\nu_1 = 0$ or 1, then $g(\nu_1; \sigma, \Delta) = 2\sigma^{-2}\left(\frac{1+\sigma}{2} - \Delta\right)^3 - 1 = f(\sigma, \Delta)$. Note also that $g(\nu_1; \sigma, \Delta)$ is continous in $\nu_1$ with a continuous first derivative $\partial g/\partial v_1$.By a similar argument to that above, the derivative of $6\nu_1(1-\nu_1) - 1 + 2\sigma^{-2}\left(\frac{1+\sigma}{2} - v_1 - \Delta\right)^3$, and hence $\partial g/\partial v_1$, is positive for $0 \leq \nu_1 \leq \frac{1+\sigma}{2} - \Delta$.

If $\sigma/2 \leq \Delta$, then by a similar argument to that above, the $\partial g/\partial v_1$ is positive for $0 \leq \nu_1 \leq \frac{1+\sigma}{2} - \Delta$. At $v_1 = 1$, $\partial g/\partial v_1 < 0$, and since for $\frac{1+\sigma}{2} - \Delta \leq \nu_1 \leq 1$ the function $g(\nu_1; \sigma, \Delta)$ is a cubic with a positive coefficient on $\nu_1{}^3$, this implies that $g(\nu_1; \sigma, \Delta)$ must reach its maximum over the domain [0, 1] for some unique $\nu_1$ between $\frac{1+\sigma}{2} - \Delta$ and 1 defined by the first order condition $\partial g/\partial v_1 = 0$.If $\sigma/2 \leq \Delta$, then $\frac{1+\sigma}{2} - \Delta \leq 0.5 \leq 1$ but $\partial g/\partial v_1 > 0$ at $v_1 = 0.5$; hence $k(\sigma, \Delta) > 0.5$. Since $\nu_1 \in \left[\frac{1+\sigma}{2} - \Delta, 1\right]$ the matching types of $\mu_1^*$ and $\mu_2^*$ are (2, 1) respectively..

If $\sigma/2 > \Delta$, then it is possible that the first order condtion is satisfied for $v_1 < \frac{1+\sigma}{2} - \Delta$.Although $\partial g/\partial v_1 > 0$ at $v_1 = 0.5$, so $k(\sigma, \Delta) > 0.5$, this would imply that the the matching types of $\mu_1^*$ and $\mu_2^*$ are (2, 2) respectively. More precisely, if $\sigma > \widetilde{\sigma}$ where $\partial g/\partial v = 0$ at $v_1 = \frac{1+\sigma}{2} - \Delta$ and $\sigma = \widetilde{\sigma}$, then $\mu_1^*$ and $\mu_2^*$ are both type 2.

Elementary calculus then shows that $\partial k(\sigma, \Delta)/\partial\Delta$ exists and is strictly negative, and that $\partial h(\sigma, \Delta)/\partial\Delta$ exists and is strictly negative, implying $h(\sigma, \Delta) > 0.5$.

In both Sections 2.1.1 and 2.1.2, $g(0; \sigma, \Delta) = g(1; \sigma, \Delta) = f(\sigma, \Delta)$, and $\partial g/\partial v_1 <$ (*resp.* $>$) 0 for $v_1 <$ (*resp.* $>$) $k(\sigma, \Delta)$, implying $g(\nu_1; \sigma, \Delta) \geq f(\sigma, \Delta)$ for all $0 < \nu_1 < 1$.

**2.2:** $\sigma = 1$

We consider the case where $0 < |\Delta| \leq 1$, with $\sigma = 1$ and $\Delta = 0$ considered in Section 3.3 of this Appendix. For $0.5 \leq |\Delta| \leq 1$, equation (20) continues to hold, and it is simple to verify that a maximum is reached when $v_1 = 0.5$ with $\mu_1$ and $\mu_2$ both being of type 1 and $h(1, \Delta) = 0.5$.

If $0 < |\Delta| < 0.5$, then equation (21) holds, and again a maximum is reached when $v_1 = 0.5$ with $\mu_1$ and $\mu_2$ both being of type 2. and $h(1, \Delta) = 0.5 + 4(0.5 - |\Delta|)^3$. This is decreasing in $|\Delta|$ .

**3.** $\mu$ **is a type 3 matching:** $|\Delta| \leq \frac{1-\sigma}{2}$

In this case, as $\nu_1$ increases from 0 to 1, the matching types of $\mu_1^*$ and $\mu_2^*$ are (1, 3) respectively, then (2, 3), then (3, 2), then (3, 1). More precisely $g(\nu_1;\sigma, \Delta) =$

$$\begin{cases} 6\nu_1(1-\nu_1) + \frac{(1+\sigma)}{2}(1-v_1)^3 - \frac{3(1-\sigma)}{2}(1-v_1)(\frac{2\Delta}{1-\sigma}+v_1)^2 & \text{if } 0 \le \nu_1 \le \frac{1-\sigma}{2} - \Delta \\ 6\nu_1(1-\nu_1) + \frac{(1+\sigma)}{2}(1-v_1)^3 - \frac{3(1-\sigma)}{2}(1-v_1)(\frac{2\Delta}{1-\sigma}+v_1)^2 + 2\sigma^{-2}(\nu_1 - \frac{1-\sigma}{2} + \Delta)^3 & \text{if } \frac{1-\sigma}{2} - \Delta \le \nu_1 \le \frac{1}{2} - \frac{\Delta}{1-\sigma} \\ 6\nu_1(1-\nu_1) + \frac{1+\sigma}{2}v_1^3 - \frac{3(1-\sigma)}{2}v_1(1-v_1-\frac{2\Delta}{1-\sigma})^2 + 2\sigma^{-2}\left(\frac{1+\sigma}{2} - \Delta - v_1\right)^3 & \text{if } \frac{1}{2} - \frac{\Delta}{1-\sigma} \le \nu_1 \le \frac{1+\sigma}{2} - \Delta \\ 6\nu_1(1-\nu_1) + \frac{1+\sigma}{2}v_1^3 - \frac{3(1-\sigma)}{2}v_1(1-v_1-\frac{2\Delta}{1-\sigma})^2 & \text{if } \frac{1+\sigma}{2} - \Delta \le \nu_1 \le 1 \end{cases}$$

If $\mu$ is a type 3 matching, the function $g$ exhibits a form of symmetry that greatly simplifies the analysis. Manipulation of the formulae above shows that for $v_1 \le \widehat{v} = \frac{1}{2} - \frac{\Delta}{1-\sigma}$

$$g(1-\lambda\nu_1;\sigma,\Delta) - f(\sigma,\Delta) = \lambda^3\{g(\nu_1;\sigma,\Delta) - f(\sigma,\Delta)\} \tag{22}$$

where $\lambda = \frac{1-\widehat{v}}{\widehat{v}}$. This has three main implications:

(i) Typically the function $g$ has two local maxima, labelled $v_1(1)$ and $v_1(2)$ such that $v_1(1) < \widehat{v} < v_1(2)$, in which case $\frac{v_1(1)}{\widehat{v}} = \frac{1-v_1(2)}{1-\widehat{v}}$ . For $\Delta > 0$ (and thus $\lambda > 1$) then for any $\nu_1 < \widehat{v}$, whatever the gain to stratification $g - f$, there exists $\nu_2 = \lambda\nu_1$ that multiplies the gain to stratification by a factor of $\lambda^3$; a global maximum will therefore be achieved by some $v_1$ greater than $\widehat{v}$ (i.e. by some $v_2$ less than $1-\widehat{v}$). Similarly if $\Delta < 0$, and $\lambda < 1$ then a global maximum will be achieved by some $v_1$ less than $v_1$ less than $\widehat{v}$

(ii) As before it is easily verified that $g(0;\sigma,\Delta) = g(1;\sigma,\Delta) = f(\sigma,\Delta) = \frac{1+\sigma}{2} - \frac{6}{1-\sigma}\Delta^2$, but in addition if $v_1 = \widehat{v}$ then $g(\nu_1;\sigma,\Delta) = f(\sigma,\Delta)$. As pointed out in Section 3, with an unstratified type 3 matching the population nevertheless "self-stratifies", in the sense that in equilibrium a matched couple are either both shorter or both taller than a height $h = \overline{w} - s_w\frac{\Delta}{1-\sigma}$. Imposing a stratification $v_1 = \widehat{v}$ would define $h$ as the boundary between the two strata and therefore have no effect.

(iii) If $\Delta = 0$ (and so $\lambda = 1$), the function $g$ is symmetric around $v_1 = 0.5$. If a global maximum is achieved by some $v_1$ it is also achieved by $1 - v_1$; this in turn implies a discontinuity in the function $k(\sigma,\Delta)$ at $\Delta = 0$, although the correspondence $K(\sigma,\Delta)$ is upper-hemi continuous.

In the light of the above, we now assume that $\Delta \ge 0$, and focus on the poperties of $g$ for $\frac{1}{2} - \frac{\Delta}{1-\sigma} \le \nu_1 \le 1$.

**3.1.** $\Delta > 0$. **and** $\sigma < 1$.

It is more convenient to express $g$ as a function $\widetilde{g}$ of $v_2$ rather than $v_1$. Then $g(\nu_1;\sigma,\Delta) = g(1-\nu_2;\sigma,\Delta) = \widetilde{g}(\nu_2;\sigma,\Delta) =$

$$\begin{cases} f(\sigma,\Delta)+\frac{3(1-\sigma)}{2}v_2(2(1-\widehat{v})-v_2)^2-\frac{3+\sigma}{2}{v_2}^3 & \text{if.} 0\leq\nu_2\leq(1-\sigma)(1-\widehat{v}) \\ f(\sigma,\Delta)+\frac{3(1-\sigma)}{2}v_2(2(1-\widehat{v})-v_2)^2-\frac{3+\sigma}{2}{v_2}^3+\frac{2}{\sigma^2}(v_2-(1-\widehat{v})(1-\sigma))^3 & \\ & \text{if } (1-\sigma)(1-\widehat{v})\leq\nu_2\leq 1-\widehat{v} \end{cases}$$

At $v_2 = 0, \widetilde{g}(\nu_2;\sigma,\Delta) = f(\sigma,\Delta)$, and $\partial\widetilde{g}/\partial v_2 = 6(1-\sigma)(1-\widehat{v})^2$. Between 0 and $(1-\sigma)(1-\widehat{v})$, it is easily shown that $\widetilde{g}$ is concave in $v_2$. At $v_2 = (1-\sigma)(1-\widehat{v})$, $\partial\widetilde{g}/\partial v_2$ $= 6(1-\sigma)(1-\widehat{v})^2(\sigma^2+\sigma-1)$, which is positive if $\sigma > \sigma^* = \frac{\sqrt{5}-1}{2}$. At $v_2 = 1-\widehat{v}$, $\widetilde{g}$ has a local minimum of $f(\sigma,\Delta)$ with (left hand derivatives) $\partial\widetilde{g}/\partial v_2 = 0$ and $\partial^2\widetilde{g}/\left(\partial v_2\right)^2 > 0$. Between $(1-\sigma)(1-\widehat{v})$ and $1-\widehat{v}$, $\widetilde{g}$ is a cubic with a positive coefficient on ${v_2}^3$, and $\partial\widetilde{g}/\partial v_2$ is continuous at $v_2 = (1-\sigma)(1-\widehat{v})$. Since $\widetilde{g}(0;\sigma,\Delta) = \widetilde{g}(1-\widehat{v};\sigma,\Delta) = f(\sigma,\Delta)$, this implies that for $0 \leq \nu_2 \leq 1-\widehat{v}$, $\widetilde{g}(\nu_2;\sigma,\Delta)$ has a maximum uniquely defined by the first order condition $\partial\widetilde{g}/\partial v_2 = 0$. If $\sigma < \sigma^*$, the maximising value of $v_2$, $1-k(\sigma,\Delta)$, is less than $(1-\sigma)(1-\widehat{v})$ and $\mu_1^*$ and $\mu_2^*$ are matching types 3 and 1 respectively. If $\sigma > \sigma^*$, $1-k(\sigma,\Delta) > (1-\sigma)(1-\widehat{v})$ and $\mu_1^*$ and $\mu_2^*$ are types 3 and 2 respectively.

It is straightforward to verify that $\partial g/dv_1 > 0$ at $v_1 = 0.5$ , so $k(\sigma,\Delta) > 0.5$.

Consider now the effects on $k(\sigma,\Delta)$ of a marginal increase in $\Delta$ when $\Delta > 0$. We can exploit the fact that over the interval $[0,1]$ the function $g(\nu_1;\sigma,\Delta)$ has two local maxima, at $v_1(1)$ and $v_1(2)$. Comparative static analysis shows that $\frac{\partial v_1(1)}{\partial\Delta}$ has the same sign as

$$\frac{\Delta}{1-\sigma}-\frac{1}{2}+v_1(1)+\frac{\psi}{\sigma^2}(v_1(1)-\frac{1-\sigma}{2}+\Delta) \tag{23}$$

where $\psi = 1$ if $(1-\sigma)\widehat{v} \leq \nu_1 \leq \widehat{v}$ and 0 otherwise. Using $\widehat{v} = \frac{1}{2}-\frac{\Delta}{1-\sigma}$ and $\frac{v_1(1)}{\widehat{v}} = \frac{1-v_1(2)}{1-\widehat{v}}$, this is readily shown to equal

$$\frac{\widehat{v}}{1-\widehat{v}}\left[\frac{1}{2}-\frac{\Delta}{1-\sigma}-v_1(2)+\frac{\psi}{\sigma^2}(\frac{1+\sigma}{2}-\Delta)-v_1(2)\right]$$

which in turn has the same sign as $\frac{\partial v_1(2)}{\partial\Delta}$. Since $\partial\widehat{v}/\partial\Delta < 0$ and $\frac{v_1(1)}{\widehat{v}} = \frac{1-v_1(2)}{1-\widehat{v}}$, they cannot both increase or remain unchanged as $\Delta$ increases. Thus $\partial k/\partial\Delta < 0$.

### 3.2. $\Delta = 0$. and $\sigma < 1$.

The analysis of 3.1 continues to apply, in that the function $\widetilde{g}(\nu_2;\sigma,0)$ is well defined and between $\nu_2 = 0$ and $\nu_2 = 0.5$ has a single maximum. There is no discontinuity at $\Delta = 0$ in $\widetilde{g}(\nu_2;\sigma,\Delta)$ or the derivative $\partial\widetilde{g}/\partial\nu_2$, so the solution to the first order condition $\partial\widetilde{g}/\partial\nu_2 = 0$ is a continuous function of $\Delta$. We denote $Lim_{\Delta\to 0}k(\sigma,\Delta)$ by $\widehat{k}(\sigma)$. We now investigate the properties of $\widehat{k}(\sigma)$ as $\sigma$ approaches 0 or 1.For $\Delta = 0$, and so $\widehat{\nu} = 0.5$, $\widetilde{g}(\nu_2;\sigma,0) =$

$$f(\sigma,0)+\frac{3(1-\sigma)}{2}v_2(1-v_2)^2-\frac{3+\sigma}{2}{v_2}^3 \qquad \text{if.} 0\leq\nu_2\leq\frac{(1-\sigma)}{2}$$

$$f(\sigma,0)+\frac{3(1-\sigma)}{2}v_2(1-v_2)^2-\frac{3+\sigma}{2}{v_2}^3+\frac{2}{\sigma^2}(v_2-\frac{(1-\sigma)}{2})^3 \text{ if } \frac{(1-\sigma)}{2}\leq \nu_2 \leq 0.5$$

For low values of $\sigma$, and in particular as $\sigma \to 0$, the condition $\partial\widetilde{g}/\partial\nu_2 = 0$ is satisfied when $\mu_1$ and $\mu_2$ are types 3 and 1 respectively, the borderline between this and a matching with types (3, 2) being defined by $\partial\widetilde{g}/\partial\nu_2 = 0$ and $\nu_2 = \frac{(1-\sigma)}{2}$, yielding the critical value of $\sigma^* = \frac{\sqrt{5}-1}{2}$. It is straightforward to show that $d\widehat{k}(\sigma)/d\sigma > 0$ for any $\sigma$ strictly between 0 and 1, since over this range $\widetilde{g}(\nu_2;\sigma,\,0)$ is strictly concave in $\nu_2$. Furthermore extending the domain of $\widetilde{g}(\nu_2;\sigma,\,0)$ to $\sigma = 0$, although economically meaningless, would not generate a discontinuity in the maximising value of $v_2$ since there would be no discontinuity in $\widetilde{g}(\nu_2;\sigma,\,0)$ or its derivatives, in particular the maximising value would be unique as $\partial^2\widetilde{g}(\nu_2;\sigma,\,0)/\left(\partial\nu_2\right)^2$ would be strictly negative at $\sigma = 0$. $Lim_{\sigma\to 0}\widehat{k}(\sigma)$ can thus be found by setting $\sigma = 0$ and solving $\partial\widetilde{g}/\partial\nu_2 = 0$, yielding a value of $\frac{3}{4}$.

By contrast, $\partial^2\widetilde{g}(\nu_2;\sigma,\,0)/\left(\partial\nu_2\right)^2 = 0$ if $\sigma = 1$. $Lim_{\sigma\to 1}\widehat{k}(\sigma)$ can be found by solving $\partial\widetilde{g}/\partial\nu_2 = 0$ for $\sigma < 1$, and taking the limit as $\sigma \to 1$. For $\sigma > \sigma^*$, $\partial\widetilde{g}/\partial\nu_2 = \frac{3(1-\sigma)}{2\sigma^2}\varphi(v_2)$ where $\varphi(v_2) = \left[4\left(1+\sigma+\sigma^2\right)v_2{}^2 - 4\left(1+\sigma^2\right)v_2 + 1 - \sigma + \sigma^2\right]$. For $\sigma < 1$, any $v_2$ is a solution of $\partial\widetilde{g}/\partial\nu_2 = 0$ if and only if it is a solution of $\varphi(v_2) = 0$. As $\sigma \to 1$, $\varphi(v_2) \to \left(6\nu_2 - 1\right)\left(2v_2 - 1\right)$, which has solutions $v_2 = 0.5$ and $v_2 = \frac{1}{6}$. The former is clearly a minimum of $\widetilde{g}(\nu_2;\sigma,\,0)$, and thus $v_2 = \frac{1}{6}$ is the interior maximum analysed above; i.e. $Lim_{\sigma\to 1}\widehat{k}(\sigma) = \frac{5}{6}$.

### 3.3 $\Delta = 0$ and $\sigma = 1$.

In this case, $\widetilde{g}(\nu_2;\sigma,\,\Delta) = f(1,\,0) = 1$ for all $v_2$.Because the sets $M$ and $W$ are identical., whatever the stratification, $M_1$ and $W_1$ are identical, as are $M_2$ and $W_2$. Thus there is perfect assortment and $h(1,\,0) = 1$. However, these are the only circumstances under which perfect assortment can be obtained, even with an optimally chosen stratification. To see this, note that if either $\Delta \neq 0$ or $\sigma \neq 1$ or both, then it cannot be the case that both $M_1 = W_1$ and $M_2 = W_2$, whatever the stratification. Therefore either $f(\sigma,\,\Delta_1)$ or $f(\sigma,\,\Delta_2)$ or both are less than 1; there is some negative assortment in one or both of the strata and hence $\widetilde{\rho} = g(\nu_1;\sigma,\,\Delta) < 1$ for any $v_1$.

Turning now to the effects on $k(\sigma,\,\Delta)$ of a marginal increase in $\Delta$, we confine ourselves to establishing that for $\Delta = 0$, $\frac{\partial h}{\partial\Delta} > 0$. It is sufficient to evaluate the derivative $\frac{\partial g}{\partial\Delta}$ at r $\Delta = 0$ and $v_1 = k(\sigma,\,0)$. Then $\frac{\partial g}{\partial\Delta}$ has the same sign as $4v_1(1 - v_1) - \psi\left(\frac{1+\sigma-2v_1}{\sigma}\right)^2$. For $\psi = 0$ this is positive. For $\psi = 1$, this is positive if $\varkappa = 2v - v^2(1+\sigma) + \sigma v - \frac{\sigma+3}{4} > 0$. Now, for $\sigma = 1$, $\varkappa = (2v-1)(1-v)$, which is positive. But $\partial\varkappa/\partial\sigma = -\left(v - 0.5\right)^2 < 0$, implying that $\frac{\partial g}{\partial\Delta} > 0$ for $\sigma < 1$.

## REFERENCES


Aumann, R. (1990), Foreward to Roth, A and Sotomayor, , T*wo sided matching, a study in game theoretic modeling and analysis Cambridge.* Cambridge, C.U.P.

Becker, G. (1981), *A Treatise on the Family,* Cambridge Ma., Harvard University Press.

Benabou, R. (1993), "Workings of a City: Location, Education, and Production", *Quarterly Journal of Economics,*108, 619-652

Clark, S. (2002), "Uniqueness of Equilibrium in Two-Sided Matching," Working Paper, Department of Economics, Edinburgh University.

Coles, M. and K. Burdett (1999), "Long-Term Partnership Formation: Marriage and Employment," *Economic Journal,* 109.

Durlauf, S. (1996a), "Neighbourhood Feedbacks, Endogenous Stratification, and Income Inequality", in Barnett, W., G. Gandalfo, and C. Hillinger, (eds.) *Dynamic Disequilibrium Modelling: Proceedings of the Ninth International Symposium on Economic Theory and Econometrics,* Cambridge, Cambridge University Press.

Durlauf, S. (1996b), "Associational Redistribution: A Defense", *Politics and Society*, 24, 391-410.

Durlauf, S. and A. Seshadri (2001), "Is Assortative Matching Efficient?", Working Paper, Department of Economics, University of Wisconsin at Madison.

Gale, D. and L. Shapley, (1962), "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9-15.

Gusfield, D. and R. Irving, (1989), *The Stable Marriage Problem: Structure and Algorithms,* Cambridge Ma., MIT Press.

Irving, R. and P. Leather, (1986) "The Complexity of Counting Stable Marriages" *SIAM Journal of Computing*, 15, 655-667.

Kiyotaki and Wright, (1989), "On Money as a Medium of Exchange", *Journal of Political Economy*, 97, 927-954.

Kremer, M. and E. Maskin, (1996), "Wage Inequality and Segregation by Skill", Working Paper 5718,, National Bureau of Economic Research.

Legros. P. and A.Newman, (2002),"Monotone Matching in Perfect and Imperfect Worlds," *Review of Ecomomic Studies*, 69, 925-942.

Legros. P. and A.Newman, (2003,) "Beauty is a Beast, Frog is a Prince: Assortative Matching with Nontransferabilities", Mimeo, ULB.

McAfee, R. (2002), "Coarse Matching," *Econometrica, 70, 2025-2034.*

Roth, A (1984), "The Evolution of the Labor Market for Medical Interns and Residents: a Case Study in Game Theory," *Journal of Political Economy*, 92, 991-1016

Roth, A and Sotomayor, (1990), T*wo sided matching, a study in game theoretic modeling and analysis Cambridge*. Cambridge, C.U.P.

Topkis, D.,(1998), *Supermodularity and Complementarity*, Princeton, New Jersey. Princeton University Press..